

Álgebra lineal

Una introducción moderna

Tercera edición

David Poole

Álgebra lineal

Una introducción moderna



Álgebra lineal

Una introducción moderna

Tercera edición

David Poole
Trent University

Traducción:
Víctor Campos Olguín
Traductor profesional

Revisión técnica:
Dr. Ernesto Filio López
Unidad Profesional Interdisciplinaria en
Ingeniería y Tecnología Avanzadas
Instituto Politécnico Nacional

M. en C. Manuel Robles Bernal
Escuela Superior de Física y Matemáticas
Instituto Politécnico Nacional



CENGAGE Learning™

Australia • Brasil • Corea • España • Estados Unidos • Japón • México • Reino Unido • Singapur

CENGAGE
Learning™

Álgebra lineal,
Una introducción moderna
Tercera Edición
David Poole

Director de producto y desarrollo Latinoamérica:
Daniel Oti Yvonnet

Director editorial y de producción Latinoamérica:
Raúl D. Zendejas Espejel

Editor:
Sergio R. Cervantes González

Coordinadora de producción editorial:
Abril Vega Orozco

Editor de producción:
Timoteo Eliosa García

Coordinador de producción:
Rafael Pérez González

Diseño de portada:
Rogelio Raymundo Reyna Reynoso

Imagen de portada:
Dreamstime

Composición tipográfica:
Rogelio Raymundo Reyna Reynoso

Corporativo Santa Fe
Av. Santa Fe núm. 505, piso 12
Col. Cruz Manca, Santa Fe
C.P. 05349, México, D.F.

Traducido del libro
Linear Algebra, A Modern Introduction 3rd ed.
David Poole
Publicado en inglés por Brooks&Cole/Cengage Learning, © 2011
ISBN-13: 978-0-538-73545-2
ISBN-10: 0-538-73545-7

Datos para catalogación bibliográfica:
Poole, David
Álgebra lineal, Una introducción moderna
Tercera Edición

ISBN-13: 978-607-481-748-5
ISBN-10: 607-481-748-0

Visite nuestro sitio en:
http://latinoamerica.cengage.com

Impreso en México
1 2 3 4 5 6 7 14 13 12 11

Para mis profesores y mis alumnos:
He aprendido mucho de cada uno
de ustedes.

Contenido

Capítulo 3 Matrices 142

Capítulo 4 Eigenvalores y eigenvectores 264

Capítulo 5 Ortogonalidad 377

Capítulo 6 Espacios vectoriales 445

Capítulo 7 Distancia y aproximación 552

Prefacio

Lo último que uno sabe cuando escribe un libro,
es qué poner primero.

—Blaise Pascal
Pensées, 1670

La tercera edición de *Álgebra lineal: una introducción moderna* conserva el enfoque y características que los usuarios encontraron como fortalezas de la edición anterior. Sin embargo, agregué nuevo material para hacer el libro útil para una audiencia más amplia y también refresqué los ejercicios.

Quiero que los alumnos vean al álgebra lineal como una materia excitante y que aprecien su tremenda utilidad. Al mismo tiempo, quiero ayudarles a dominar los conceptos y técnicas básicas del álgebra lineal que necesitarán en otros cursos, tanto en matemáticas como en otras disciplinas. También quiero que los alumnos aprecien la interacción de las matemáticas teóricas, aplicadas y numéricas que impregna la materia.

Este libro está diseñado para usarse en un curso introductorio de álgebra lineal en uno o dos semestres. Primero, y más importante, está dirigido a los alumnos, e hice mi mejor esfuerzo para escribir el libro de modo que ellos no sólo lo encuentren legible, sino también que *quieran* leerlo. Como en la primera y segunda ediciones, tomé en cuenta la realidad de que los alumnos que cursan álgebra lineal introductoria probablemente provengan de varias disciplinas. Además de las especialidades en matemáticas, es adecuado para los especializados en ingeniería, física, química, ciencias de la computación, biología, ciencias ambientales, geografía, economía, psicología, negocios y educación, así como otros alumnos que toman el curso como una materia optativa o para cumplir los requisitos del grado. En concordancia, el libro equilibra teoría y aplicaciones, está escrito en un estilo conversacional aunque es completamente riguroso, y combina una presentación tradicional con la preocupación por el aprendizaje centrado en el alumno.

No existe como un mejor estilo de aprendizaje único. En cualquier grupo habrá algunos alumnos que trabajen bien de manera independiente y otros que trabajen mejor en grupos; algunos que prefieran el aprendizaje basado en conferencias y otros que prosperen en un escenario de talleres, realizando exploraciones; algunos que gocen con las manipulaciones algebraicas, algunos que sean fanáticos del cálculo numérico (con y sin computadora) y algunos que muestren fuerte intuición geométrica. En este libro continúo la presentación del material en varias formas (*algebraica, geométrica, numérica* y *verbal*) de modo que todos los tipos de aprendices puedan encontrar una ruta a seguir. También traté de presentar los temas teórico, de cálculo y aplicación en una forma flexible, aunque integrada. Al hacerlo, espero que todos los alumnos estarán expuestos a los muchos lados del álgebra lineal.

Vea *The College Mathematics Journal* **24** (1993), 41-46.

Este libro es compatible con las recomendaciones del Lineal Algebra Curriculum Study Group. Desde un punto de vista pedagógico, no hay duda de que, para la mayoría de los

alumnos, los ejemplos concretos deben anteceder a la abstracción. Aquí tomé este enfoque. También creo firmemente que el álgebra lineal trata esencialmente de vectores y que los alumnos deben ver primero vectores (en un escenario concreto) con la finalidad de obtener cierta comprensión geométrica. Más aún, la introducción temprana de los vectores permite a los alumnos ver cómo los sistemas de ecuaciones lineales surgen de manera natural de problemas geométricos. Luego las matrices surgen igualmente de manera natural como matrices de coeficientes de sistemas lineales y como agentes de cambio (transformaciones lineales). Esto prepara el escenario para los eigenvectores y las proyecciones ortogonales, los cuales se entienden mejor geométricamente. Los dardos que aparecen en la cubierta de este libro simbolizan vectores y reflejan mi convicción de que la comprensión geométrica debe anteceder a las técnicas de cálculo.

Traté de limitar el número de teoremas en el texto. En su mayor parte, los resultados marcados como teoremas se usarán más tarde en el texto o resumen el trabajo precedente. Los resultados interesantes que no son el tema central del libro se incluyeron como ejercicios o exploraciones. Por ejemplo, el producto cruz de vectores se estudia sólo en exploraciones (en los capítulos 1 y 4). A diferencia de la mayoría de los libros de álgebra lineal, el que tiene en sus manos no tiene un capítulo acerca de determinantes. Los resultados esenciales están todos en la sección 4.2, con otro material interesante contenido en una exploración. Sin embargo, el libro es muy completo para ser un texto introductorio. Siempre que fue posible, incluí pruebas elementales y accesibles a los teoremas, con la finalidad de evitar el tener que decir: "la prueba de este resultado está más allá del ámbito de este texto". El resultado es, espero, una obra completa en sí misma.

No fui tacaño con las aplicaciones: existen muchas más en el libro de las que se pueden abarcar en un solo curso. Sin embargo, es importante que los alumnos vean la impresionante gama de problemas a la que puede aplicarse el álgebra lineal. Incluí algún material moderno acera de teoría de codificación que por lo general no se encuentra en un texto introductorio de álgebra lineal. También hay impresionantes aplicaciones del álgebra lineal en el mundo real, y un tema de interés histórico, si no práctico, que se presenta como "viñetas" independientes.

Espero que los profesores disfruten la enseñanza a partir de este libro. Más importante: espero que los alumnos que usen el libro terminen el curso apreciando la belleza, el poder y la tremenda utilidad del álgebra lineal y que tengan diversión a lo largo del camino.

Qué hay de nuevo en la tercera edición

La estructura global y el estilo de *Álgebra lineal: una introducción moderna* siguen siendo los mismos que en la tercera edición.

He aquí un resumen del material nuevo:

Vea las páginas 13, 50

- El capítulo 1 se reorganizó de manera extensa. La introducción a la aritmética modular y al álgebra lineal finita se movieron a la sección 1.1. La sección 1.4 ahora sólo contiene aplicaciones: vectores código y una nueva subsección acerca de vectores fuerza, como aparecen en física e ingeniería.

Vea las páginas 113, 241

- Los modelos económicos lineales, un tema de importancia en negocios y economía, se agregaron como aplicación en los capítulos 2 y 3.

Vea la página 295

- Una nueva viñeta acerca del "método de condensación" de Lewis Carroll para evaluar determinantes se agregó al capítulo 4.
- Hay más de 400 ejercicios nuevos o revisados.
- Realicé numerosos cambios pequeños de redacción para mejorar la claridad o precisión de la exposición.

Vea la página xvi

- Las asignaciones de tareas en línea ahora pueden realizarse usando *Enhanced WebAssign*, que contiene ejercicios relacionados con el libro.

- Ejercicios, con soluciones, ahora están disponibles para los Apéndices en el sitio web para el libro a disposición del profesor. Esto facilitará el uso del libro para aquellos profesores que incluyan material de los apéndices en sus cursos.
- Todos los auxiliares existentes se actualizaron y muchos se colocaron en el sitio web de la empresa.

Vea la página xv

Características

Estilo de escritura claro

El texto está escrito en un estilo simple, directo y conversacional. Tanto como fue posible, usé "idioma matemático" en lugar de apoyarme de manera excesiva en notación matemática. Sin embargo, todas las pruebas dadas son totalmente rigurosas y el Apéndice A contiene una introducción a la notación matemática para quienes deseen simplificar su propia escritura. Ejemplos concretos casi siempre preceden a los teoremas, a los que luego siguen más ejemplos y aplicaciones. Este flujo, de lo específico a lo general y de regreso, es consistente a lo largo del libro.

Conceptos clave introducidos al inicio

Muchos alumnos encuentran dificultad en el álgebra lineal cuando el curso avanza de la realización de cálculos (resolver sistemas de ecuaciones lineales, manipulación de vectores y matrices) a lo teórico (espacios generadores, independencia lineal, subespacios, bases y dimensión). Este libro introduce todos los conceptos clave del álgebra lineal muy temprano, en un escenario concreto, antes de volver a visitarlos con total generalidad. Los conceptos vectoriales como producto punto, longitud, ortogonalidad y proyección se estudian primero en el capítulo 1, en el escenario concreto de $\mathbb{R}^2$ y $\mathbb{R}^3$ antes de que las nociones más generales de producto interior, norma y proyección ortogonal aparezcan en los capítulos 5 y 7. De igual modo, a los espacios generadores y la independencia lineal se les da un tratamiento concreto en el capítulo 2, antes de su generalización a espacios vectoriales en el capítulo 6. Los conceptos fundamentales de subespacio, base y dimensión aparecen primero en el capítulo 3, cuando se introducen los espacios renglón, columna y nulo de una matriz; no es sino hasta el capítulo 6 cuando a estas ideas se les da un tratamiento general. En el capítulo 4 se introducen los eigenvalores y los eigenvectores, y se exploran para matrices de 2×2 antes de que aparezcan sus contrapartes de $n \times n$. Al comenzar el capítulo 4, todos los conceptos clave del álgebra lineal se han introducido, con ejemplos de cálculo concretos para apoyarlos. Cuando dichas ideas aparecen con plena generalidad más tarde en el libro, los alumnos han tenido tiempo de acostumbrarse a ellos y, por tanto, no se sienten intimidados.

Énfasis en vectores y geometría

Al conservar la filosofía de que el álgebra lineal trata principalmente con vectores, este libro subraya la intuición geométrica. En concordancia, el primer capítulo es acerca de vectores y desarrolla muchos conceptos que aparecerán de manera repetida a lo largo del texto. Conceptos como ortogonalidad, proyección y combinación lineal se encuentran todos en el capítulo 1, así como un tratamiento amplio de líneas y planos en $\mathbb{R}^3$ que ofrece comprensión esencial a la solución de sistemas de ecuaciones lineales. Este énfasis en vectores, geometría y visualización se encuentra a lo largo del texto. Las transformaciones lineales se introducen como transformaciones de matrices en el capítulo 3, con muchos ejemplos geométricos, antes de incluir las transformaciones lineales generales en el capítulo 6. En el capítulo 4 se introducen los eigenvalores con "eigenimágenes"

como auxiliar visual. La prueba del Teorema de Perron se proporciona primero heurísticamente y luego formalmente, en ambos casos usando un argumento geométrico. La geometría de los sistemas dinámicos lineales refuerza y resume el material acerca de eigenvalores y eigenvectores. En el capítulo 5, las proyecciones ortogonales, los complementos ortogonales de subespacios y el proceso Gram-Schmidt se presentan todos en el escenario concreto de $\mathbb{R}^3$ antes de generalizarse a $\mathbb{R}^n$ y, en el capítulo 7, a espacios de producto interior. La naturaleza de la descomposición del valor singular también se explica de manera informal en el capítulo 7 vía un argumento geométrico. De las más de 300 figuras en el texto, más de 200 se dedican a fomentar una comprensión geométrica del álgebra lineal.

Exploraciones

Vea las páginas 1, 142, 445, 552

Vea las páginas 32, 297, 478, 533, 566, 570

Vea las páginas 89, 90, 91, 407, 409

La introducción a cada capítulo es una exploración guiada (sección 0) en la que se invita a los alumnos a descubrir, individualmente o en grupos, algún aspecto del capítulo venidero. Por ejemplo, "El juego de la pista de carreras" introduce los vectores, "Matrices en acción" introduce la multiplicación de matrices y las transformaciones lineales, "Fibonacci en el espacio (vectorial)" toca conceptos de espacio vectorial y "Geometría de taxi" establece normas generalizadas y funciones de distancia. Las exploraciones adicionales que se encuentran a lo largo del libro incluyen aplicaciones de vectores y determinantes a la geometría, una investigación de los cuadrados mágicos de 3×3, un estudio de la simetría basado en los mosaicos de M. C. Escher, una introducción al álgebra lineal compleja y problemas de optimización usando desigualdades geométricas. También existen exploraciones que introducen importantes consideraciones numéricas y el análisis de algoritmos. Hacer que los alumnos realicen algunas de estas exploraciones es una forma de alentarlos a convertirse en aprendices activos y de darles "dominio" sobre una pequeña parte del curso.

Aplicaciones

Vea las páginas 57, 545

Vea las páginas 60, 127, 232, 367, 630

El libro contiene una abundante selección de aplicaciones elegidas de una amplia gama de disciplinas, incluidas matemáticas, ciencias de la computación, física, química, ingeniería, biología, negocios, economía, psicología, geografía y sociología. Digno de atención entre éstas es un fuerte tratamiento de la teoría de codificación, desde los códigos de detección de error (como el International Standard Book Numbers) hasta sofisticados códigos de corrección de error (como el código Reed-Muller, que se usó para transmitir fotografías satelitales desde el espacio). Adicionalmente, existen cinco "viñetas" que presentan brevemente algunas aplicaciones muy modernas del álgebra lineal: el sistema Codabar, el sistema de posicionamiento global (GPS), robótica, motores de búsqueda en Internet y la compresión de imágenes digitales.

Ejemplos y ejercicios

Existen más de 400 ejemplos en este libro, la mayoría resueltos con mayor detalle de lo acostumbrado en un libro de texto introductorio al álgebra lineal. Este nivel de detalle es consecuente con la filosofía de que los alumnos querrán (y podrán) leer un libro de texto. En concordancia, no se tiene la intención de que todos los ejemplos se estudien en clase; muchos pueden asignarse para estudio individual o grupal, posiblemente como parte de un proyecto. La mayoría de los ejemplos tienen al menos un ejercicio contraparte, de modo que los alumnos pueden practicar las habilidades incluidas en el ejemplo antes de explorar las generalizaciones.

Hay más de 2000 ejercicios, más que en la mayoría de los libros de texto de nivel similar. La respuesta a la mayoría de los ejercicios de cálculo con número impar puede en-

contrarse en la parte final del libro. Los profesores descubrirán una riqueza de ejercicios de dónde elegir para asignaciones de tarea en casa. (En la *Guía del Instructor* en línea se brindan algunas sugerencias.) Los ejercicios en cada sección están graduados y avanzan de lo rutinario a lo desafiante. Los ejercicios varían desde los que tienen la intención del cálculo mental a los que requieren el uso de una calculadora o un sistema algebraico de cómputo, y de los ejercicios teóricos y numéricos a los ejercicios conceptuales. Muchos de los ejemplos y ejercicios usan datos reales recopilados de situaciones del mundo real. Por ejemplo, existen problemas acerca de modelado del crecimiento de poblaciones de caribúes y focas, datación con radiocarbono del monumento de Stonehenge y predicción de salarios de jugadores de las grandes ligas de beisbol. Trabajar tales problemas refuerza el hecho de que el álgebra lineal es una valiosa herramienta para el modelado de problemas de la vida real.

Vea las páginas 257, 370, 547, 611

Ejercicios adicionales aparecen en forma de repaso al final de cada capítulo. En cada conjunto hay 10 preguntas verdadero/falso diseñadas para poner a prueba la comprensión conceptual, seguidas por 19 ejercicios de cálculo y teóricos que resumen los principales conceptos y técnicas de cada capítulo.

Bosquejos biográficos y notas etimológicas

Es importante que los alumnos aprendan algo acerca de la historia de las matemáticas y lleguen a verla como un esfuerzo social y cultural, así como científico. En concordancia, el texto contiene breves bosquejos biográficos acerca de muchos de los matemáticos que contribuyeron al desarrollo del álgebra lineal. Espero que estos bosquejos ayuden a poner un rostro humano a la materia y que proporcionen a los alumnos otra forma de relacionarse con el material. La *Guía del Instructor* en línea sugiere formas para ampliar algunas de tales notas biográficas en proyectos de ensayo.

Descubrí que muchos alumnos se sienten alejados de las matemáticas porque la terminología no tiene sentido para ellos: se trata simplemente de una colección de palabras por aprender. Para ayudar a superar este problema, incluí breves notas etimológicas que ofrecen los orígenes de muchos de los términos que se usan en álgebra lineal. (Por ejemplo, ¿por qué se usa la palabra *normal* para referirse a un vector que es perpendicular a un plano?)

Vea la página 34

Iconos marginales

Los márgenes del libro contienen muchos iconos cuyo propósito es alertar al lector en varias formas. El cálculo no es un prerrequisito para este libro, pero el álgebra lineal tiene muchas aplicaciones interesantes e importantes para el cálculo. El icono $\frac{dy}{dx}$ denota un ejemplo o ejercicio que requiere cálculo. (Este material puede omitirse si nadie en el grupo ha cursado al menos un semestre de cálculo. Alternativamente, pueden asignarse como proyectos.) El icono $a + bi$ denota un ejemplo o ejercicio que involucra números complejos. (Para los alumnos no familiarizados con los números complejos, el Apéndice C contiene todo el material de antecedentes que se necesita.) El icono CAS indica que se requiere, o al menos es muy útil, un sistema algebraico de cómputo (como Maple, Mathematica o MATLAB) o una calculadora con capacidades para matrices (como la TI-89, TI-Nspire, HP-48gII, HP-50g, Casio 9850GC+ o Sharp EL-9200C) para resolver el ejemplo o ejercicio.

Con la intención de ayudar a los alumnos a aprender cómo leer y usar este libro de texto de manera más efectiva, anoté varios lugares donde se aconseja al lector detenerse por un momento. Se trata de lugares donde se necesita un cálculo, debe suministrarse parte de una prueba, debe verificarse una afirmación o se requiere algo de pensamiento adicional. El icono → aparece en el margen de tales lugares; el mensaje es: "Deténgase. Saque su lápiz. Piense en esto".

Tecnología

Este libro puede usarse exitosamente ya sea que los alumnos tengan o no acceso a la tecnología. Sin embargo, las calculadoras con capacidades para matrices y los sistemas algebraicos de cómputo son ahora comunes y, utilizados de manera adecuada, pueden enriquecer la experiencia de aprendizaje así como ayudar con los cálculos tediosos. En este texto se toma el punto de vista de que los alumnos necesitan dominar todas las técnicas básicas del álgebra lineal al resolver ejemplos a mano que no son demasiado difíciles para realizar cálculos. Entonces la tecnología puede usarse (en todo o en parte) para resolver ejemplos y aplicaciones posteriores, y para aplicar técnicas que se apoyen en los primeros. Por ejemplo, cuando se introducen por primera vez sistemas de ecuaciones lineales, se proporcionan soluciones detalladas; más tarde, simplemente se presentan las soluciones, y se espera que el lector las compruebe. Este es un buen lugar para usar alguna forma de tecnología. Del mismo modo, cuando las aplicaciones usen datos que hagan imprácticos los cálculos a mano, use tecnología. Todos los métodos numéricos que se estudian dependen del uso de tecnología.

Con la ayuda de tecnología, los alumnos pueden explorar el álgebra lineal en ciertas formas excitantes y descubrir mucho por ellos mismos. Por ejemplo, si uno de los coeficientes de un sistema lineal se sustituye por un parámetro, ¿cuánta variabilidad hay en las soluciones? ¿Cómo cambiar una sola entrada de una matriz afecta sus eigenvalores? Este libro no es un curso acerca de tecnología, y en los lugares donde ésta puede usarse, no se especifica un tipo particular de tecnología. El sitio web para el alumno que acompaña a este libro ofrece un apéndice en línea llamado *Technology Bytes* que ofrece instrucciones para resolver una selección de ejemplos de cada capítulo usando Maple, Mathematica y MATLAB. Al imitar dichos ejemplos, los alumnos pueden realizar más cálculos y exploraciones usando cualquier CAS (siglas en inglés de sistema algebraico de cómputo), MATLAB que tengan y explotar el poder de dichos sistemas para auxiliarse con los ejercicios a lo largo del libro, en particular con los marcados con el icono CAS. El sitio web también contiene conjuntos de datos y código de cómputo en formatos Maple, Mathematica y MATLAB relacionados con muchos ejercicios y ejemplos en el texto. Alumnos y profesores pueden importarlos directamente en su CAS para conservar la escritura y eliminar errores.

Álgebra lineal finita y numérica

El texto abarca dos aspectos del álgebra lineal que escasamente se mencionan juntos: álgebra lineal finita y álgebra lineal numérica. Al introducir pronto aritmética modular, es posible hacer del álgebra lineal finita (más correctamente, "álgebra lineal sobre campos finitos", aunque yo no uso dicha frase) un tema recurrente a lo largo del libro. Este enfoque brinda acceso al material acerca de teoría de codificación en las secciones 1.4, 3.7, 5.5, 6.7 y 7.5. También hay una aplicación a juegos lineales finitos en la sección 2.4 que los alumnos realmente gozarán. Además de estar expuestos a las aplicaciones de álgebra lineal finita, quienes se especialicen en matemáticas se beneficiarán de ver el material acerca de campos finitos, porque probablemente los encontrarán en cursos como matemáticas discretas, álgebra abstracta y teoría de números.

Vea las páginas 53, 111, 251, 419, 543, 640

Todos los alumnos deben tener en cuenta que, en la práctica, en el álgebra lineal es imposible llegar a soluciones exactas de problemas de gran escala. La exposición a algunas de las técnicas del álgebra lineal numérica brindará un indicio de cómo obtener soluciones aproximadas enormemente precisas. Algunos de los temas numéricos incluidos en el libro son: errores de redondeo y pivoteo parcial, métodos iterativos para resolver sistemas lineales y eigenvalores de computación, las factorizaciones *LU* y *QR*, las normas matriciales y números de condición, aproximación por mínimos cuadrados y el valor singular de descomposición. La inclusión del álgebra lineal numérica también plantea algunos problemas interesantes e importantes que están ausentes por completo de la *teo-*

Vea las páginas 89, 90, 130, 322, 392, 403, 584, 591, 613

ría del álgebra lineal, como las estrategias de pivoteo, la condición de un sistema lineal y la convergencia de los métodos iterativos. Este libro no sólo plantea dichas cuestiones, sino también muestra cómo puede abordarlos. Los discos de Gerschgorin, las normas matriciales y los valores singulares de una matriz, que se estudian en los capítulos 4 y 7, son útiles a este respecto.

Vea las páginas 330, 586, 623

Apéndices

El Apéndice A contiene un panorama de la notación matemática y métodos de prueba, y el Apéndice B discute la inducción matemática. Todos los alumnos se beneficiarán de dichas secciones, pero quienes estén orientados a una especialidad en matemáticas acaso deban poner particular atención en ellos. Algunos de los ejemplos en dichos apéndices son raros (por ejemplo, el ejemplo B.6 del Apéndice B) y subrayan el poder de los métodos. El Apéndice C es una introducción a los números complejos. Para alumnos familiarizados con dichos resultados, este apéndice puede servir como una útil referencia; para otros, esta sección contiene todo lo que necesitan conocer para aquellas partes del texto que usan números complejos. El Apéndice D es acerca de polinomios. Descubrí que muchos alumnos requieren repasar estos temas. La mayoría de los alumnos no está familiarizado con la regla de los signos de Descartes; se usa en el capítulo 4 para explicar el comportamiento de los eigenvalores de las matrices de Leslie. Los ejercicios que acompañan los cuatro apéndices pueden encontrarse en el sitio web del libro.

Al final del libro se proporcionan respuestas cortas a la mayoría de los ejercicios de cálculo con número impar. Los conjuntos de ejercicios para acompañar los Apéndices A, B, C y D están disponibles en el sitio web acompañante, junto con sus respuestas de número impar.

Auxiliares

Los siguientes complementos están todos disponibles de manera gratuita para los profesores que adopten *Álgebra lineal: una introducción moderna* (tercera edición). Los alumnos pueden comprar el *Manual de Soluciones y Guía de Estudio del Alumno*, por separado o adjunto con el libro. El sitio web tiene secciones protegidas con contraseña para alumnos y profesores.

Manual de Soluciones y Guía de Estudio del Alumno
ISBN-10: 0-538-73771-9;
ISBN-13: 978-0-538-73771-5
Incluye soluciones detalladas de todos los ejercicios con número impar y de ejercicios seleccionados con número par; resúmenes de sección y capítulo de símbolos, definiciones y teoremas; y consejos y sugerencias de estudio. Los ejercicios complejos se exploran mediante un formato pregunta/respuesta diseñado para profundizar en el conocimiento. También se incluyen problemas desafiantes y de entretenimiento que exploran aún más ejercicios seleccionados.

Solution Builder
www.cengage.com/solutionbuilder
Ofrece al profesor soluciones completas de todos los ejercicios del texto, incluidos los que se encuentran en las exploraciones y repasos de capítulo, en conveniente formato en línea. Solution Builder permite a los profesores crear impresiones personalizadas seguras de soluciones en PDF que coinciden exactamente con los ejercicios asignados al grupo. Disponible para quienes adopten el libro al firmar en la dirección web mencionada anteriormente.

Guía del Instructor
www.cengage.com/math/poole, sitio web para el profesor
Esta guía en línea, escrita por Douglas Shaw y Michael Prophet, aumenta el texto con valiosos recursos pedagógicos como proyectos de trabajo grupal, consejos pedagógicos, interesantes preguntas de examen, ejemplos y material adicional para clases, y otros ítems diseñados para reducir el tiempo de preparación del profesor y hacer que las clases de álgebra lineal sean una experiencia atractiva e interactiva. Para cada sección del texto, la *Guía del Instructor* incluye tiempo y énfasis sugeridos, puntos a resaltar, preguntas para discusión, materiales y ejemplos de clase, consejos tecnológicos, proyectos estudiantiles, trabajo grupal con soluciones, ejemplos de tareas y sugerencias de preguntas para examen.

ExamView® banco de exámenes electrónico
ISBN-10: 0-538-73770-0; ISBN-13: 978-0-538-73770-8
El software para exámenes ExamView permite a los profesores crear, entregar y personalizar rápidamente exámenes para sus grupos en formatos impresos y en línea, y características de calificación automática. Incluye más de 500 preguntas verdadero/falso, opción múltiple y respuesta libre con base en el texto. Todos los ítems de examen también se proporcionan en formatos PDF y Microsoft® Word para los profesores que opten por no usar el componente de software.

Enhanced WebAssign
Un sistema de enseñanza/aprendizaje fácil de usar en línea que ofrece tareas para casa asignables, calificación automática, asistencia interactiva para estudiantes y control de gestión de curso para profesores. Con cientos de ejercicios acordes al texto, los alumnos obtienen práctica para resolver problemas que clarifican el álgebra lineal, construyen habilidades e impulsan la comprensión conceptual. La interfase simple y amigable con el usuario de *Enhanced WebAssign* permite a los profesores crear rápidamente un curso, inscribir alumnos, seleccionar problemas para una tarea y controlar el número de intentos de respuesta que se permitan a los alumnos. Una boleta de calificaciones rica en características ayuda a administrar las calificaciones del grupo, establecer curvas de calificación, asignar fechas límite y exportar resultados a una hoja de cálculo fuera de línea. Para más información, visite www.webassign.net/brookscole.

Sitio web para Álgebra lineal: Una introducción moderna
www.cengage.com/math/poole
Contiene materiales adicionales en línea para acompañar el texto dirigidos a alumnos y profesores. Aquí pueden encontrarse ejercicios para acompañar los apéndices del libro, junto con respuestas seleccionadas. El apéndice *Technology Bytes* ofrece instrucciones CAS para Maple, Mathematica y MATLAB para resolver ejemplos de cada capítulo. Conjuntos de datos CAS descargables en formatos Maple, Mathematica y MATLAB ofrecen código directo de computadora para trabajar ejercicios y ejemplos del texto en un CAS. La *Guía del Instructor*, versiones estáticas de los ítems de examen *ExamView* y otros recursos útiles también son accesibles aquí sólo para los profesores.

Reconocimientos

Los revisores de cada edición de este texto aportaron valiosos y con frecuencia perspicaces comentarios acerca del libro. Estoy agradecido por el tiempo que cada uno de ellos tomó para hacer esto. Su juicio y útiles sugerencias contribuyeron enormemente al desarrollo y éxito de este libro, y quiero agradecerles personalmente:
Mark Brittenham, University of Nebraska; Marek Elzanowski, Portland State University; J. Douglas Faires, Youngstown State University; Yuval Flicker, The Ohio State University; Justin Travis Gray, Simon Fraser University; William Hager, University of

Florida; Sasho Kalajdzievski, University of Manitoba; Israel Koltracht, University of Connecticut; Dr. En-Bing Lin, University of Toledo; Dr. Asamoah Nkwanta, Morgan State University; Gleb Novitchkov, Penn State University; Arthur Robinson, George Washington University; Roman Smirnov, Dalhousie University; Ron Solomon, Ohio State University; Mo Tavakoli, Chaffey College; Bruno Welfert, Arizona State University.

Estoy en deuda con mucha gente que, a lo largo de los años, influyó en mis percepciones acerca del álgebra lineal y la enseñanza de las matemáticas en general. Primero, quiero agradecer de manera colectiva a los participantes en las sesiones de educación y álgebra lineal especial en reuniones de la Mathematical Association of America y la Canadian Mathematical Society. También aprendí mucho de la participación en la Canadian Mathematics Education Study Group y el Canadian Mathematics Education Forum.

Quiero agradecer especialmente a Ed Barbeau, Bill Higginson, Richard Hoshino, John Grant McLoughlin, Eric Muller, Morris Orzech, Bill Ralph, Pat Rogers, Peter Taylor y Walter Whiteley, cuyos consejos e inspiración contribuyeron enormemente con la filosofía y el estilo de este libro. Mi gratitud también para Robert Rogers, quien desarrolló las soluciones para alumnos y profesores, así como el excelente contenido de la guía de estudio. Un agradecimiento especial para Jim Stewart por su apoyo y consejos continuos. Joe Rotman y su adorable libro *A First Course in Abstract Algebra* inspiró las notas etimológicas en este libro, y me apoyé bastante en el *The Words of Mathematics* de Steven Schwartzman cuando compilé dichas notas. Agradezco a Art Benjamin por introducirme en el sisetma Codabar. Mis colegas Marcus Pivato y Reem Yassawi ofrecieron información útil acerca de los sistemas dinámicos. Como siempre, estoy agradecido con mis alumnos por plantear buenas preguntas y proporcionarme la realimentación necesaria para convertirme en un mejor profesor.

Sinceramente agradezco a todas las personas que estuvieron involucradas en la producción de este libro. Gunjan Chandola y su equipo en MPS Limited realizaron una sorprendente labor al producir la tercera edición con un calendario muy apretado. Agradezco a Roger Lipsett por realizar una excelente labor al verificar la precisión de esta edición y a Richard Camp por hacer una concienzuda corrección del libro. Ha sido una experiencia deliciosa trabajar con los autores de los auxiliares, quienes estuvieron en la misma longitud de onda que yo desde el principio. Doug Shaw y Mike Prophet escribieron una excelente *Guía del Instructor*; Richard Pappas expandió y mejoró el contenido del banco de exámenes disponible en ExamView. Sobre todo, ha sido una delicia trabajar con todos los equipos editoriales, de marketing y de producción en Cengage Learning: Molly Taylor, editora sponsor; Dan Seibert, editor asociado; Shaylin Walsh, asistente editorial; Susan Miscio, gerente de proyecto de contenido; Andrew Coppola, editor de medios; y Jennifer Jones, gerente de marketing. Ellos ofrecieron atinados consejos acerca de cambios y adiciones, ofrecieron ayuda cuando la necesitaba, pero me permitieron escribir el libro que quería. Soy afortunado por haber trabajado con ellos, así como con el personal de las ediciones primera y segunda.

Como siempre, agradezco a mi familia por su amor, apoyo y comprensión. Sin ellos, este libro no habría sido posible.

David Poole
dpoole@trentu.ca

Para el profesor

"¿Me dirías, por favor,
hacia dónde debo ir desde aquí?"
"Eso depende bastante
de hacia dónde quieras ir", dijo el Gato.

—Lewis Carroll
Alice's Adventures in
Wonderland, 1865

Este texto se escribió con flexibilidad en mente. Tiene la intención de usarse en un curso de uno o dos semestres con 36 clases por semestre. La gama de temas y aplicaciones lo hace adecuado para varias audiencias y tipos de cursos. Sin embargo, hay más material en el libro del que puede estudiarse en clase, incluso en un curso de dos semestres. Después del siguiente panorama del texto hay algunas breves sugerencias acerca de formas de usar el libro. La *Guía del Instructor* en línea tiene sugerencias más detalladas, incluidas notas pedagógicas, ejercicios recomendados, actividades y proyectos para el salón de clase, y temas adicionales

Panorama del texto

Capítulo 1: Vectores

Vea la página 1

El juego de la pista de carreras en la sección 1.0 sirve para introducir los vectores de manera informal. (¡También es muy divertido para jugar!) Luego los vectores se introducen formalmente tanto desde un punto de vista algebraico como de uno geométrico. Las operaciones de suma y multiplicación escalar y sus propiedades se desarrollan primero en los escenarios concretos de $\mathbb{R}^2$ y $\mathbb{R}^3$ antes de generalizarse a $\mathbb{R}^n$. También se introducen la aritmética modular y el álgebra lineal finita. La sección 1.2 define el producto punto de vectores y las nociones relacionadas de longitud, ángulo y ortogonalidad. El concepto muy importante de proyección (ortogonal) se desarrolla aquí; reaparecerá en los capítulos 5 y 7. La exploración "Vectores y geometría" muestra cómo pueden usarse métodos vectoriales para probar ciertos resultados en la geometría euclideana. La sección 1.3 es una introducción básica, aunque amplia, a las líneas y los planos en $\mathbb{R}^2$ y $\mathbb{R}^3$. Esta sección es crucial para comprender el significado geométrico de la solución de los sistemas lineales del capítulo 2. Note que el producto cruz de vectores en $\mathbb{R}^3$ se deja como exploración. El capítulo concluye con dos aplicaciones: vectores fuerza y vectores código. La mayoría de los alumnos disfrutarán la aplicación al Universal Product Code (UPC) y al International Standard Book Number (ISBN). La viñeta acerca del sistema Codabar que se usa en tarjetas de crédito y bancarias es una excelente presentación para clase que incluso puede usarse para introducir la sección 1.4

Vea la página 32

Vea la página 48

Vea las páginas 56, 57, 60

Capítulo 2: Sistemas de ecuaciones lineales

Vea la página 63

Vea las páginas 78, 211, 397, 504

Vea la página 127

Vea las páginas 89, 90, 91

La introducción a este capítulo sirve para ilustrar que existe más de una forma de encontrar una solución para un sistema de ecuaciones lineales. Las secciones 2.1 y 2.2 desarrollan la principal herramienta de evaluación para resolver sistemas lineales: reducción de matrices (eliminación gaussiana y de Gauss-Jordan). Casi todos los métodos de evaluación posteriores en el libro dependen de esto. El teorema del rango aparece aquí por primera vez; se muestra de nuevo, con más generalidad, en los capítulos 3, 5, y 6. La sección 2.3 es muy importante: ahí se introducen las nociones fundamentales de generadores y de independencia lineal de vectores. No se apresure con este material. La sección 2.4 contiene seis aplicaciones que pueden elegir los profesores, dependiendo del tiempo disponible y de los intereses del grupo. La viñeta acerca del Sistema de Posicionamiento Global ofrece otra aplicación que disfrutarán los alumnos. Los métodos iterativos de la sección 2.5 serán opcionales para muchos cursos, pero son esenciales para un curso con un enfoque aplicado/numérico. Las tres exploraciones de este capítulo se relacionan en que todos tratan con aspectos del uso de las computadoras para resolver sistemas lineales. Todos los alumnos deben al menos estar al tanto de estos temas.

Capítulo 3: Matrices

Vea la página 142

Vea las páginas 178, 212, 307, 530, 628

Vea la página 232

Vea las páginas 236, 245

Este capítulo contiene algunas de las ideas más importantes en el libro. Es un capítulo largo, pero el material inicial puede estudiarse muy rápidamente, con tiempo adicional reservado para el material crucial de la sección 3.5. La sección 3.0 es una exploración que introduce la noción de transformación lineal: la idea de que las matrices no sólo son objetos estáticos sino más bien un tipo de función que transforma vectores en otros vectores. Todos los hechos básicos acerca de las matrices, las operaciones matriciales y sus propiedades se encuentran en las primeras dos secciones. El material acerca de matrices particionadas y las representaciones múltiples del producto matricial valen la pena destacarse, porque se usan repetidamente en secciones posteriores. El teorema fundamental de las matrices invertibles de la sección 3.3 es muy importante y aparecerá muchas veces más conforme se presenten nuevas caracterizaciones de invertibilidad. La sección 3.4 discute la muy importante factorización *LU* de una matriz. Si este tema no se estudia en clase, vale la pena asignarlo como proyecto o discutirlo en un taller. El punto de la sección 3.5 es presentar muchos de los conceptos clave del álgebra lineal (subespacio, base, dimensión y rango) en el escenario concreto de matrices antes de que los alumnos los vean con plena generalidad. Aunque los ejemplos de esta sección son todos familiares, es importante que los alumnos se acostumbren a la nueva terminología y, en particular, entiendan qué significa la noción de base. El tratamiento geométrico de las transformaciones lineales de la sección 3.6 tiene la intención de suavizar la transición hacia las transformaciones lineales generales del capítulo 6. El ejemplo de una proyección es particularmente importante porque reaparecerá en el capítulo 5. La viñeta acerca de brazos robóticos es una demostración concreta de la composición de transformaciones lineales (y afines). Existen cinco aplicaciones de las cuales elegir en la sección 3.7. Deben estudiarse ya sean las cadenas de Markov o el modelo de Leslie de crecimiento poblacional, de modo que puedan usarse nuevamente en el capítulo 4, donde se explicará su comportamiento.

Capítulo 4: Eigenvalores y eigenvectores

Vea la página 264

La introducción a la sección 4.0 presenta un interesante sistema dinámico que involucra gráficas. Esta exploración introduce la noción de un eigenvector y anuncia el método de potencias en la sección 4.5. Para conservar el énfasis geométrico del libro, la sección 4.1 contiene la novedosa característica de “eigenimágenes” como una forma de visualizar los eigenvectores de matrices de 2×2. Los determinantes aparecen en la sección 4.2, moti-

vados por su uso en el descubrimiento de los polinomios característicos de matrices pequeñas. Este "curso intensivo" de determinantes contiene todo el material esencial que necesitan los alumnos, incluida una prueba opcional pero elemental del teorema de expansión de Laplace. La viñeta "Método de condensación de Lewis Carroll" presenta un método alternativo, con interés histórico, para calcular determinantes que los alumnos encontrarán muy atractivo. La exploración "Aplicaciones geométricas de los determinantes" constituye un buen proyecto que contiene muchos resultados interesantes y útiles. (Alternativamente, los profesores que deseen poner más atención en los determinantes pueden elegir exponer parte de esta exploración en clase.) La teoría básica de los eigenvalores y los eigenvectores se encuentra en la sección 4.3, y la sección 4.4 trata el importante tema de la diagonalización. El ejemplo 4.29 acerca de potencias de matrices vale la pena estudiarlo en clase. El método de potencias y sus variantes, que se estudian en la sección 4.5, son opcionales, pero todos los alumnos deben estar al tanto del método, y un curso aplicado debe analizarlo con detalle. El teorema del disco de Gerschgorin puede estudiarse independientemente del resto de la sección 4.6. Aunque la prueba del Teorema de Perron es opcional, el teorema en sí (como el más fuerte teorema de Perron-Frobenius) debe al menos mencionarse, porque explica *por qué* debemos esperar un eigenvalor positivo único con un correspondiente eigenvector positivo en dichas aplicaciones. Las aplicaciones acerca de relaciones de recurrencia y ecuaciones diferenciales conectan el álgebra lineal con las matemáticas discretas y el cálculo, respectivamente. La exponencial de matrices puede exponerse si su grupo tiene un buen antecedente de cálculo. El tema final de los sistemas dinámicos lineales discretos vuelve a revisar y resume muchas de las ideas del capítulo 4, y las observa bajo una nueva luz geométrica. Los alumnos gozarán leyendo cómo pueden usarse los eigenvectores para ayudar a clasificar equipos deportivos y sitios web. Esta viñeta puede ampliarse fácilmente a un proyecto o actividad enriquecedora.

Vea la página 295

Vea la página 297

Vea las páginas 336, 341

Vea la página 367

Capítulo 5: Ortogonalidad

La exploración introductoria, "Sombras en la pared", es la mejor matemática: toma un concepto conocido (proyección de un vector sobre otro vector) y lo generaliza en una forma útil (proyección de un vector sobre un subespacio: un plano), mientras descubre algunas propiedades anteriormente no observadas. La sección 5.1 contiene los resultados básicos acerca de conjuntos de vectores ortogonales y ortonormales que se usarán repetidamente de aquí en adelante. En particular, deben destacarse las matrices ortogonales. En la sección 5.2 se generalizan dos conceptos del capítulo 1: el complemento ortogonal de un subespacio y la proyección ortogonal de un vector sobre un subespacio. El teorema de descomposición ortogonal es importanteaquí y ayuda a configurar el proceso Gram-Schmidt. Note también la rápida prueba del teorema del rango. El proceso Gram-Schmidt se detalla en la sección 5.3, junto con la extremadamente importante factorización *QR*. Las dos exploraciones que siguen subrayan cómo se calcula en la práctica la factorización *QR* y cómo puede usarse para aproximar eigenvalores. La sección 5.4 acerca de diagonalización ortogonal de matrices simétricas (reales) es necesaria para las aplicaciones que continúan. También contiene el teorema espectral, uno de los puntos destacados de la teoría del álgebra lineal. Las aplicaciones de la sección 5.5 incluyen códigos duales, formas cuadráticas y graficación de ecuaciones cuadráticas. Yo siempre incluyo al menos la última de éstas en mi curso, porque amplía lo que ya conocen los alumnos acerca de las secciones cónicas.

Vea la página 377

Vea las páginas 407, 409

Vea las páginas 419, 425, 432

Capítulo 6: Espacios vectoriales

La secuencia de Fibonacci reaparece en la sección 6.0, aunque no es importante que los alumnos la hayan visto antes (sección 4.6). El propósito de esta exploración es mostrar que los conceptos familiares de espacio vectorial (sección 3.5) pueden usarse fructífera-

Vea la página 445

mente en un nuevo escenario. Puesto que todas las ideas principales de los espacios vectoriales ya se introdujeron en los capítulos 1-3, los alumnos encontrarán las secciones 6.1 y 6.2 bastante familiares. El énfasis aquí debe ser sobre el uso de los axiomas de espacio vectorial para probar propiedades en lugar de apoyarse en técnicas para calcular. Cuando se discute el cambio de base en la sección 6.3, es útil mostrar a los alumnos cómo usar la notación para recordar cómo funciona la construcción. A final de cuentas, el método Gauss-Jordan es el más eficiente aquí. Son importantes las secciones 6.4 y 6.5 acerca de transformaciones lineales. Los ejemplos se relacionan con resultados anteriores acerca de matrices (y de transformaciones matriciales). En particular, es importante subrayar que el kernel y el rango de una transformación lineal generalizan el espacio nulo y el espacio de columna de una matriz. La sección 6.6 profundiza en la noción de que (casi) todas las transformaciones lineales son en esencia transformaciones matriciales. Esto se apoya en la información de la sección 3.6, de modo que los alumnos no deben encontrarla terriblemente sorprendente. Sin embargo, los ejemplos deben resolverse cuidadosamente. La conexión entre cambio de base y similitud de matrices es digna de atención. La exploración “Mosaicos, retículas y la restricción cristalográfica” es una impresionante aplicación del cambio de base. La conexión con la obra de M. C. Escher la hace todavía más interesante. Las aplicaciones de la sección 6.7 se apoyan en las anteriores y pueden incluirse según lo permitan el tiempo y el interés.

Vea la página 533

Capítulo 7: Distancia y aproximación

Vea la página 552

Vea la página 566

Vea la página 570

La sección 7.0 abre con la entretenida exploración “Geometría de taxi”. Su propósito es establecer el material acerca de las funciones norma y distancia generalizadas (métricas) que se presentan después. Los espacios de producto interior se estudian en la sección 7.1; el énfasis aquí debe estar en los ejemplos y usar los axiomas. La exploración “Vectores y matrices con entradas complejas” muestra cómo pueden extenderse los conceptos de producto punto, matriz simétrica, matriz ortogonal y diagonalización ortogonal, desde los espacios vectoriales reales hasta los complejos. La siguiente exploración, “Desigualdades geométricas y problemas de optimización”, es una que usualmente disfrutan los alumnos. (¡Se divertirán al ver cuántos problemas de “cálculo” pueden resolverse sin usar cálculo!) La sección 7.2 incluye normas vectoriales y matriciales generalizadas y muestra cómo el número condicional de una matriz se relaciona con la noción de sistemas lineales mal condicionados que se exploró en el capítulo 2. La aproximación de mínimos cuadrados (sección 7.3) es una importante aplicación del álgebra lineal en muchas otras disciplinas. El teorema de mejor aproximación y el teorema de mínimos cuadrados son importantes, pero sus pruebas son intuitivamente claras. Dedique algo de tiempo a los ejemplos, algunos serán suficientes. La sección 7.4 presenta la descomposición en valor singular, una de las aplicaciones más importantes del álgebra lineal. Si su curso llega hasta aquí, será ampliamente recompensado. La SVD no sólo vincula muchas nociones discutidas anteriormente; también ofrece algunas aplicaciones novedosas (y bastante poderosas). Si está disponible un CAS, la viñeta acerca de compresión de imagen digital vale la pena de presentar; es un impresionante despliegue visual del poder del álgebra lineal y una culminación que ajusta bien con el curso. Las posteriores aplicaciones de la sección 7.5 pueden elegirse de acuerdo con el tiempo disponible y el interés del grupo.

Vea la página 624

Cómo usar el libro

Los alumnos encontrarán el libro muy fácil de leer, por lo que yo generalmente les dejo leer una sección antes de exponer el material en clase. De esa forma puedo emplear el tiempo de clase destacando los conceptos más importantes y lidiar con los temas que a los alumnos les parecen difíciles, para trabajar ejemplos y para discutir aplicaciones. No

trato de abarcar todo el material de la lectura asignada en clase. Este enfoque me permite mantener el ritmo del curso bastante vivaz y freno un poco para aquellas secciones que usualmente a los alumnos les resultan desafiantes.

En un curso de dos semestres es posible revisar todo el libro, incluida una razonable selección de aplicaciones. Para flexibilidad adicional, puede omitir algunos de los temas (por ejemplo, ofrecer sólo un breve tratamiento del álgebra lineal numérica), de ese modo libero tiempo para un análisis más a profundidad de los temas restantes, más aplicaciones o alguna de las exploraciones. En un curso de matemáticas decente que enfatice las pruebas, gran parte del material en los capítulos 1-3 puede estudiarse rápidamente. Entonces el capítulo 6 puede exponerse en conjunción con las secciones 3.5 y 3.6, y el capítulo 7 puede integrarse en el capítulo 5. Yo me aseguraría de asignar las exploraciones de los capítulos 1, 4, 6 y 7 para tales clases.

Para un curso de un semestre, la naturaleza del curso y la audiencia determinarán cuáles temas incluir. A continuación se describen tres posibles cursos. El curso básico, descrito primero, tiene menos de 36 horas sugeridas, lo que permite tiempo para temas adicionales, revisión en clase y exámenes. Los otros dos cursos se apoyan en el curso básico, pero todavía son bastante flexibles.

Un curso básico

A continuación se bosqueja un curso diseñado para especialidades en matemáticas y alumnos de otras disciplinas. Este curso no menciona espacios vectoriales generales en absoluto (todos los conceptos se tratan en un escenario concreto) y es muy ligero en las pruebas. Sin embargo, es una introducción completa al álgebra lineal.

Sección	Número de clases	Sección	Número de clases
1.1	1	**3.6**	1–2
1.2	1–1.5	**4.1**	1
1.3	1–1.5	**4.2**	2
2.1	0.5–1	**4.3**	1
2.2	1–2	**4.4**	1–2
2.3	1–2	**5.1**	1–1.5
3.1	1–2	**5.2**	1–1.5
3.2	1	**5.3**	0.5
3.3	2	**5.4**	1
3.5	2	**7.3**	2

Total: 23–30 clases

Dado que los alumnos en un curso como este representan una amplia variedad de disciplinas, yo sugeriría el uso de gran parte del tiempo de clase restante para aplicaciones. En mi curso lo hago con los vectores de código de la sección 1.4, lo que realmente parece agradar a los alumnos, y al menos una aplicación de cada uno de los capítulos 2-5. Otras aplicaciones pueden asignarse como proyectos, junto con tantas de las exploraciones como se desee. También existe suficiente tiempo de clase disponible para estudiar parte de la teoría con detalle.

Un curso con énfasis en la realización de cálculos

Para un curso con énfasis en la realización de cálculos, el curso básico que se presentó en la página anterior puede complementarse con las secciones del texto que tratan con el álgebra lineal numérica. En tal caso, yo expondría en parte o completas las secciones 2.5, 3.4, 4.5, 5.3, 7.2 y 7.4, para terminar con la descomposición en valores singulares. Las exploraciones de los capítulos 2 y 5 son particularmente adecuadas para tal curso, como lo son casi cualquiera de las aplicaciones.

Un curso para alumnos que ya estudiaron algo de álgebra lineal

Algunos cursos estarán dirigidos a alumnos que ya encontraron los principios básicos del álgebra lineal en otros cursos. Por ejemplo, un curso de álgebra universitaria con frecuencia incluirá una introducción a los sistemas de ecuaciones lineales, matrices y determinantes; un curso de cálculo en muchas variables casi seguramente contendrá material acerca de vectores, líneas y planos. Para los alumnos que ya han visto tales temas, puede omitirse gran parte del material inicial y sustituirlo con un repaso rápido. Dependiendo de los antecedentes del grupo, puede ser posible revisar rápidamente el material en el curso básico hasta la sección 3.3 aproximadamente en seis clases. Si el grupo tiene un número significativo de especialistas en matemáticas (y especialmente si este es el único curso de álgebra lineal que tendrán), yo me aseguraría de exponer las secciones 1.4, 6.1-6.5, 7.1 y 7.4 y tantas aplicaciones como permita el tiempo. Si el curso tiene especialistas en ciencia (mas no especialistas en matemáticas), abarcaría las secciones 1.4, 6.1 y 7.1 y una selección más amplia de aplicaciones, para estar seguro de incluir el material acerca de ecuaciones diferenciales y aproximación de funciones. Si están más representados los alumnos de ciencias de la computación o ingenieros, trataría de exponer tanto material como pudiera acerca de códigos y álgebra lineal numérica.

Existen muchos otros tipos de cursos que pueden usar con éxito este libro. Espero que usted lo encuentre útil para su curso y que se divierta al usarlo.

Para el alumno

"¿Por dónde debo comenzar, por favor, su majestad?", preguntó.
"Comience por el principio", dijo el rey, con gravedad, "y continúe hasta que llegue al final: entonces deténgase."

—Lewis Carroll
Alice's Adventures in Wonderland, 1865

El álgebra lineal es una materia excitante. Está llena de resultados interesantes, aplicaciones para otras disciplinas y conexiones hacia otras áreas de las matemáticas. El *Manual de soluciones y guía de estudio del estudiante* contiene consejos detallados acerca de cómo usar mejor este libro; a continuación se ofrecen algunas sugerencias generales.

El álgebra lineal tiene varios lados: existen *técnicas para realizar cálculos*, *conceptos* y *aplicaciones*. Una de las metas de este libro es ayudarlo a dominar todas estas facetas de la materia y a ver la interacción entre ellas. En consecuencia, es importante que lea y comprenda cada sección del texto antes de abordar los ejercicios en dicha sección. Si sólo lee los ejemplos que se relacionan con los ejercicios que se asignaron como tarea para casa, perderá mucho. Asegúrese de entender las definiciones de términos y el significado de los teoremas. No se preocupe si tiene que leer algo más de una vez antes de comprenderlo. Tenga a mano lápiz y calculadora mientras lee. Deténgase para trabajar los ejemplos por usted mismo o para completar los cálculos faltantes. El icono ➞ en el margen indica un lugar donde debe detenerse y pensar en lo que ha leído hasta el momento.

Las respuestas a la mayoría de los ejercicios con número impar para realizar cálculos están al final del libro. Resista a la tentación de mirar la respuesta antes de haber completado una pregunta. Y recuerde que, incluso si su respuesta difiere de la del libro, todavía puede estar correcto; hay más de una forma correcta de expresar algunas de las soluciones. Por ejemplo, un valor de $1/\sqrt{2}$ también puede expresarse como $\sqrt{2}/2$, y el conjunto de todos los múltiplos escalares del vector $\begin{bmatrix} 3 \\ 1/2 \end{bmatrix}$ es el mismo que el conjunto de todos los múltiplos escalares de $\begin{bmatrix} 6 \\ 1 \end{bmatrix}$.

Conforme encuentre nuevos conceptos, trate de relacionarlos con los ejemplos que conozca. Escriba las pruebas y soluciones a los ejercicios en una forma lógica y conectada, y use oraciones completas. Lea de nuevo lo que escribió para ver si tiene sentido. Mejor aún: si puede, haga que un amigo del grupo lea lo que usted escribió. Si no tiene sentido para otra persona, hay posibilidades de que no tenga sentido, punto.

Descubrirá que son útiles o una calculadora con capacidades para matrices o un sistema algebraico de cómputo. Dichas herramientas pueden ayudarle a verificar sus pro-

pios cálculos a mano y son indispensables para algunos problemas que involucran cálculos tediosos. La tecnología también le permite explorar aspectos del álgebra lineal por cuenta propia. Puede jugar juegos de "¿y si…?": ¿y si cambio uno de los elementos en este vector? ¿Y si esta matriz es de un tamaño diferente? ¿Puedo forzar la solución para que sea lo que quiero que sea al cambiar algo? Para señalar los lugares en el texto o los ejercicios donde se recomienda el uso de tecnología, coloqué el icono CAS en el margen. El sitio web que acompaña a esta obra contiene ejercicios seleccionados del libro, resueltos con código de computadora usando Maple, Mathematica y MATLAB, así como *Technology Bytes*, un apéndice que proporciona múltiples consejos adicionales acerca del uso de la tecnología en el álgebra lineal.

Está a punto de embarcarse en un viaje a través del álgebra lineal. Piense en este libro como en su guía de viaje. ¿Está listo? ¡Adelante!

1 Vectores

Aquí vienen, surgiendo de la tristeza. Pequeñas flechas para mí y para ti.

—Albert Hammond y Mike Hazelwood
Little Arrows
Dutchess Music/BMI, 1968

1.0 Introducción: el juego de la pista de carreras

Muchas cantidades mensurables, como la longitud, el área, el volumen, la masa y la temperatura, pueden describirse completamente al especificar su magnitud. Otras cantidades, como la velocidad, la fuerza y la aceleración, requieren tanto una magnitud como una dirección para su descripción. Estas cantidades son *vectores*. Por ejemplo, la velocidad del viento es un vector que consta de rapidez del viento y su dirección, como 10 km/h suroeste. Geométricamente, los vectores se representan con frecuencia como flechas o segmentos de recta dirigidos.

Aunque la idea de vector se introdujo en el siglo XIX, su utilidad práctica, particularmente en las ciencias físicas, no se conoció sino hasta el siglo XX. Más recientemente, los vectores encontraron aplicaciones en ciencias de la computación, estadística, economía y las ciencias de la vida y sociales. A lo largo de este libro se considerarán algunas de estas muchas aplicaciones.

Este capítulo introduce los vectores y comienza por considerar algunas de sus propiedades geométricas y algebraicas. También se considerarán aplicaciones no geométricas en las que son útiles los vectores. Sin embargo, se comienza con un juego simple que presenta algunas de las ideas clave. [Incluso quizá quiera jugarlo con un amigo durante los momentos aburridos (¡muy raros!) en las clases de álgebra lineal.]

El juego se juega en papel gráfico. Sobre éste se dibuja una pista, con una línea de partida y una línea de meta. La pista puede ser de cualquier longitud y forma, en tanto sea suficientemente ancha para alojar a todos los jugadores. Para este ejemplo, se tendrán dos jugadores (llámelos Ann y Bert), que usan plumas de diferentes colores para representar sus carros o bicicletas o cualquier cosa que empleen para correr alrededor de la pista. (Piense en Ann y Bert como ciclistas.)

Ann y Bert comienzan cada uno por dibujar un punto en la línea de partida en un lugar de la cuadrícula sobre el papel gráfico. Toman turnos para moverse hacia un nuevo punto de la cuadrícula, sujetos a las siguientes reglas:

1. Cada nuevo punto en la cuadrícula y el segmento de recta que lo conecta con el punto anterior en la cuadrícula debe encontrarse completamente dentro de la pista.
2. Dos jugadores no pueden ocupar el mismo punto en la cuadrícula en el mismo turno. (Esta es la regla de "no colisión".)
3. Cada nuevo movimiento se relaciona con el movimiento anterior del modo siguiente: si un jugador se mueve a unidades horizontalmente y b unidades vertical-

mente en un movimiento, entonces en el siguiente movimiento debe moverse entre $a - 1$ y $a + 1$ unidades horizontalmente, y entre $b - 1$ y $b + 1$ unidades verticalmente. En otras palabras, si el segundo movimiento es c unidades horizontalmente y d unidades verticalmente, entonces $|a - c| \leq 1$ y $|b - d| \leq 1$. (Esta es la regla "aceleración/desaceleración".) Note que esta regla fuerza al primer movimiento a ser 1 unidad verticalmente y/o 1 unidad horizontalmente.

Un jugador que choque con otro o deje la pista, es eliminado. El ganador es el primer jugador en cruzar la línea de meta. Si más de un jugador cruza la línea de meta en el mismo turno, quien las sobrepase más es el ganador.

En el juego de muestra que se presenta en la figura 1.1, Ann fue la ganadora. Bert aceleró muy rápidamente y tuvo dificultades para tomar la curva en la parte superior de la pista.

Para comprender la regla 3, considere el tercer y cuarto movimientos de Ann. En su tercer movimiento, avanzó 1 unidad horizontalmente y 3 unidades verticalmente. En su cuarto movimiento, sus opciones eran moverse de 0 a 2 unidades horizontalmente y de 2 a 4 unidades verticalmente. (Note que algunas de dichas combinaciones la habrían dejado fuera de la pista.) Eligió moverse 2 unidades en cada dirección.

El matemático irlandés William Rowan Hamilton (1805–1865) usó conceptos vectoriales en su estudio de los números complejos y su generalización, los cuaterniones.

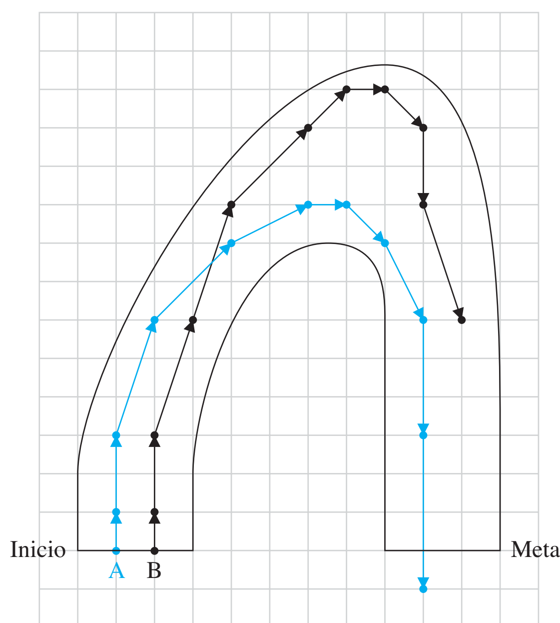


Figura 1.1
Un ejemplo del juego de pista de carrera

Problema 1 Juegue algunos juegos de la pista de carreras.

Problema 2 ¿Es posible que Bert gane esta carrera si elige una secuencia diferente de movimientos?

Problema 3 Use la notación $[a, b]$ para denotar un movimiento que sea a unidades horizontalmente y b unidades verticalmente. (Tanto a o b o ambos pueden ser negativos.) Si acaba de realizar el movimiento $[3, 4]$, dibuje en el papel gráfico todos los puntos de la cuadrícula que posiblemente podría alcanzar en el siguiente movimiento.

Problema 4 ¿Cuál es el efecto *neto* de dos movimientos sucesivos? En otras palabras, si se mueve $[a, b]$ y luego $[c, d]$, ¿cuán lejos, horizontal y verticalmente, se habrá movido?

Problema 5 Escriba la secuencia de movimientos de Ann con la notación $[a, b]$. Suponga que ella comienza en el origen (0, 0) en los ejes coordenados. Explique cómo puede encontrar las coordenadas del punto de cuadrícula correspondiente a cada uno de sus movimientos *sin mirar el papel gráfico*. Si los ejes se dibujan de manera diferente, de modo que el punto de partida de Ann no esté en el origen sino en el punto (2, 3), ¿cuáles serían las coordenadas de su punto final?

Aunque simple, este juego introduce muchas ideas que serán útiles en el estudio de los vectores. Las siguientes tres secciones consideran a los vectores desde los puntos de vista geométrico y algebraico, comenzando, como en el juego de la pista de carreras, en el plano.

1.1 Geometría y álgebra de vectores

Vectores en el plano

Comience por considerar el plano cartesiano con los familiares ejes x y y. Un ***vector*** es un *segmento de recta dirigido* que corresponde a un *desplazamiento* desde un punto A hasta otro punto B; vea la figura 1.2.

El vector de A a B se denota mediante $\overrightarrow{AB}$; el punto A se conoce como su ***punto inicial*** u ***origen***, y el punto B se conoce como su ***punto terminal*** o ***punta***. Con frecuencia, un vector simplemente se denota mediante una sola letra minúscula negrita, como $\mathbf{v}$.

El conjunto de todos los puntos en el plano corresponden al conjunto de todos los vectores cuyos orígenes están en el origen O. A cada punto A, le corresponde el vector $\mathbf{a} = \overrightarrow{OA}$; a cada vector $\mathbf{a}$ con origen en O, le corresponde su punta A. (Los vectores de esta forma en ocasiones se conocen como *vectores de posición*.)

Es natural representar dichos vectores usando coordenadas. Por ejemplo, en la figura 1.3, $A = (3, 2)$ y el vector se escribe $\mathbf{a} = \overrightarrow{OA} = [3, 2]$ con corchetes. De igual modo, los otros vectores en la figura 1.3 son

$$\mathbf{b} = [-1, 3] \quad \text{y} \quad \mathbf{c} = [2, -1]$$

Las coordenadas individuales (3 y 2 en el caso de $\mathbf{a}$) se llaman los ***componentes*** del vector. En ocasiones, se dice que un vector es un *par ordenado* de números reales. El orden es importante pues, por ejemplo, $[3, 2] \neq [2, 3]$. En general, dos vectores son iguales si y sólo si sus componentes correspondientes son iguales. Por tanto, $[x, y] = [1, 5]$ implica que $x = 1$ y $y = 5$.

Con frecuencia es conveniente usar ***vectores columna*** en lugar de (o además de) ***vectores renglón***. Otra representación de $[3, 2]$ es $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. (El punto importante es que los

El plano cartesiano recibe su nombre en honor del filósofo y matemático francés René Descartes (1596–1650), cuya introducción de las coordenadas permitió que los problemas *geométricos* se manejaran con el uso de técnicas *algebraicas*.

La palabra *vector* proviene de la raíz latina que significa "transportar". Un vector se forma cuando un punto se desplaza, o "transporta", una distancia dada en una dirección dada. Visto de otra forma, los vectores "transportan" dos piezas de información: su longitud y su dirección.

Cuando los vectores se escriben a mano, es difícil indicar las negritas. Algunas personas prefieren escribir $\vec{v}$ para el vector denotado en impreso por $\mathbf{v}$, pero en la mayoría de los casos está bien usar una minúscula ordinaria v. Por lo general será claro desde el contexto cuando la letra denota un vector.

La palabra *componente* se deriva de las palabras latinas *co*, que significa "junto con", y *ponere*, que significa "poner". Por tanto, un vector es "poner juntos" sus componentes.

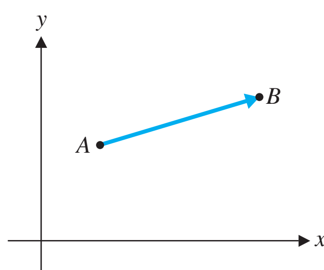


Figura 1.2

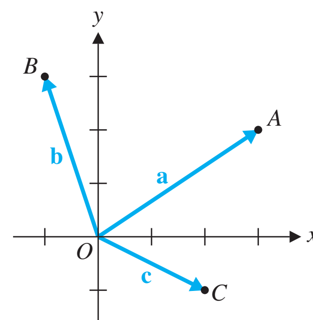


Figura 1.3

componentes están *ordenados*.) En capítulos posteriores verá que los vectores columna son un tanto mejores desde el punto de vista computacional; por ahora, trate de acostumbrarse a ambas representaciones.

Puede ocurrir que en realidad no pueda dibujar el vector $[0, 0] = \overrightarrow{OO}$ desde el origen hacia sí mismo. No obstante. Es un vector perfectamente bueno y tiene un nombre especial: el ***vector cero***. El vector cero se denota **0**.

$\mathbb{R}^2$ se pronuncia "r dos".

El conjunto de todos los vectores con dos componentes se denota $\mathbb{R}^2$ (donde $\mathbb{R}$ denota el conjunto de números reales de donde se eligen los componentes de los vectores en $\mathbb{R}^2$). Por tanto, $[-1, 3.5]$, $[\sqrt{2}, \pi]$ y $[\frac{5}{3}, 4]$ están todos en $\mathbb{R}^2$.

Piense de nuevo en el juego de la pista de carreras y trate de conectar todas estas ideas con los vectores cuyos orígenes *no* están en el origen. El origen etimológico de la palabra *vector* en el verbo "transportar" ofrece una pista. El vector [3, 2] puede interpretarse del modo siguiente: a partir del origen O, viaje 3 unidades a la derecha, luego 2 unidades arriba y termine en P. El mismo desplazamiento puede aplicarse con otros puntos iniciales. La figura 1.4 muestra dos desplazamientos equivalentes, representados por los vectores $\overrightarrow{AB}$ y $\overrightarrow{CD}$.

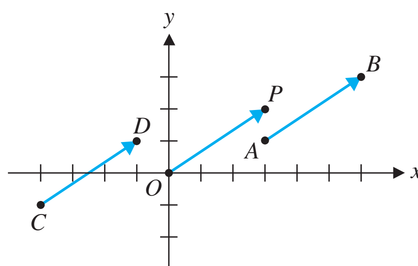


Figura 1.4

Cuando se hace referencia a los vectores mediante coordenadas, se les considera *analíticamente*.

Dos vectores se definen como *iguales* si tienen la misma longitud y la misma dirección. Por tanto, $\overrightarrow{AB} = \overrightarrow{CD}$ en la figura 1.4. (Aun cuando tengan diferentes puntos inicial y final, representan el mismo desplazamiento.) Geométricamente, dos vectores son iguales si uno puede obtenerse al deslizar (o *trasladar*) el otro paralelo a sí mismo hasta que los dos vectores coincidan. En términos de componentes, en la figura 1.4 se tiene $A = (3, 1)$ y $B = (6, 3)$. Note que el vector [3, 2] que registra el desplazamiento sólo es la diferencia de los componentes respectivos:

$$\overrightarrow{AB} = [3, 2] = [6 - 3, 3 - 1]$$

De igual modo, $$\overrightarrow{CD} = [-1 - (-4), 1 - (-1)] = [3, 2]$$

y por tanto $\overrightarrow{AB} = \overrightarrow{CD}$, como se esperaba.

Se dice que un vector como $\overrightarrow{OP}$ con su punto inicial en el origen está en ***posición estándar***. La discusión anterior muestra que todo vector puede dibujarse como un vector en posición estándar. Por otro lado, un vector en posición estándar puede redibujarse (por traslación) de modo que su origen esté en cualquier punto en el plano.

Ejemplo 1.1

Si $A = (-1, 2)$ y $B = (3, 4)$, encuentre $\overrightarrow{AB}$ y vuelva a dibujarlo (a) en posición estándar y (b) con su origen en el punto $C = (2, -1)$.

Solución Calcule $\overrightarrow{AB} = [3 - (-1), 4 - 2] = [4, 2]$. Si entonces $\overrightarrow{AB}$ se traslada hacia $\overrightarrow{CD}$, donde $C = (2, -1)$, entonces se debe tener $D = (2 + 4, -1 + 2) = (6, 1)$. (Vea la figura 1.5.)

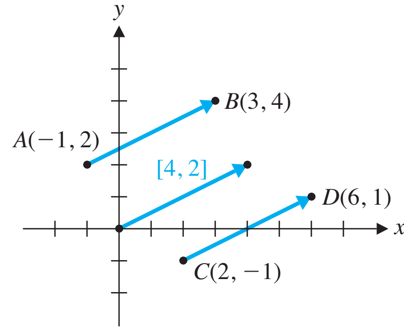
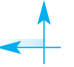


Figura 1.5



Nuevos vectores a partir de otros anteriores

Como en el juego de la pista de carreras, con frecuencia se quiere "seguir" un vector a partir de otro. Esto conduce a la noción de ***suma de vectores***, la primera operación vectorial básica.

Si se sigue **u** por **v**, se puede visualizar el desplazamiento total como un tercer vector, denotado mediante **u** + **v**. En la figura 1.6, **u** = [1, 2] y **v** = [2, 2], de modo que el efecto neto de seguir **u** por **v** es

$$[1 + 2, 2 + 2] = [3, 4]$$

lo que produce **u** + **v**. En general, si **u** = $[u_1, u_2]$ y **v** = $[v_1, v_2]$, entonces su ***suma*** **u** + **v** es el vector

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2]$$

Es útil visualizar **u** + **v** geométricamente. La siguiente regla es la versión geométrica de la discusión anterior.

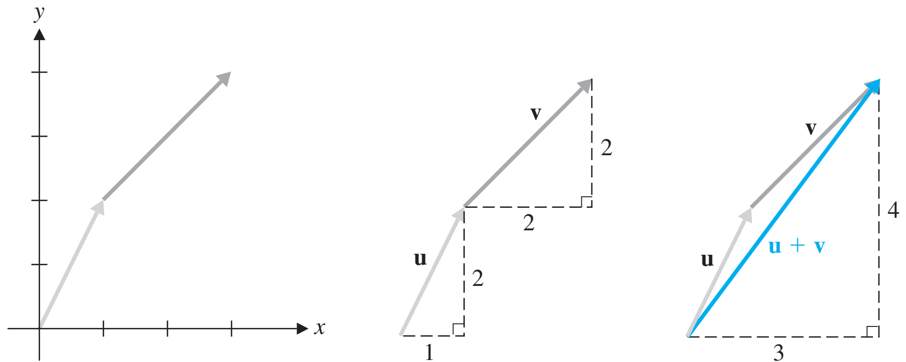


Figura 1.6
Suma de vectores

La regla punta a origen

Dados los vectores **u** y **v** en $\mathbb{R}^2$, traslade **v** de modo que su origen coincida con la punta de **u**. La ***suma*** **u** + **v** de **u** y **v** es el vector desde el origen de **u** hasta la punta de **v**. (Vea la figura 1.7.)

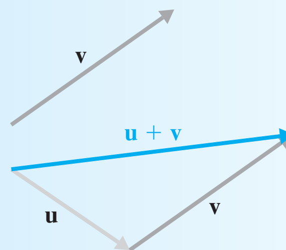


Figura 1.7
La regla punta a origen

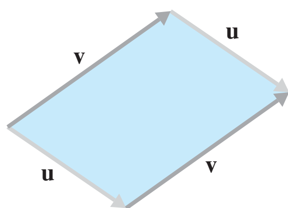


Figura 1.8
El paralelogramo determinado por **u** y **v**

Al trasladar **u** y **v** paralelos a ellos mismos, se obtiene un paralelogramo, como se muestra en la figura 1.8. Este paralelogramo se llama *paralelogramo determinado por* **u** *y* **v**. Ello conduce a una versión equivalente de la regla punta a origen para vectores en posición estándar.

La regla del paralelogramo

Dados los vectores **u** y **v** en $\mathbb{R}^2$ (en posición estándar), su ***suma*** **u** + **v** es el vector en posición estándar a lo largo de la diagonal del paralelogramo determinado por **u** y **v**. (Vea la figura 1.9.)

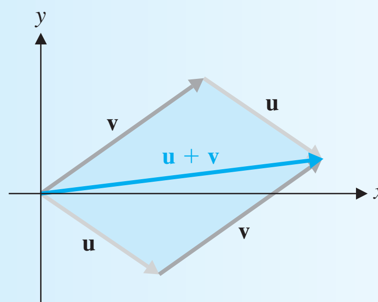


Figura 1.9
La regla del paralelogramo

Ejemplo 1.2

Si **u** = [3, −1] y **v** = [1, 4], calcule y dibuje **u** + **v**.

Solución Calcule **u** + **v** = [3 + 1, −1 + 4] = [4, 3]. Este vector se dibuja mediante la regla punta a origen en la figura 1.10(a) y mediante la regla del paralelogramo en la figura 1.10(b).

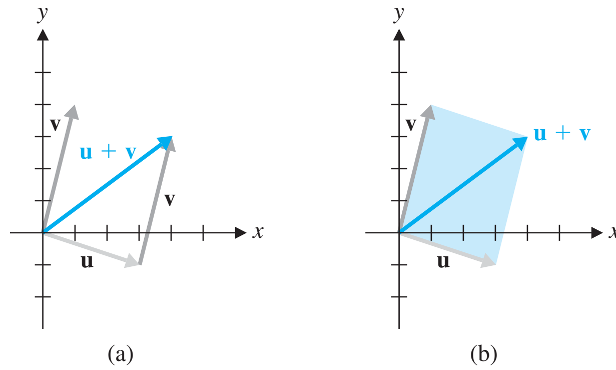


Figura 1.10

La segunda operación vectorial básica es la ***multiplicación escalar***. Dado un vector **v** y un número real c, el ***múltiplo escalar*** $c\mathbf{v}$ es el vector que se obtiene al multiplicar cada componente de **v** por c. Por ejemplo, $3[-2, 4] = [-6, 12]$. En general,

$$c\mathbf{v} = c\,[v_1, v_2] = [cv_1, cv_2]$$

Geométricamente, $c\mathbf{v}$ es una versión "a escala" de **v**.

Ejemplo 1.3

Si $\mathbf{v} = [-2, 4]$, calcule y dibuje $2\mathbf{v}$, $\frac{1}{2}\mathbf{v}$, y $-2\mathbf{v}$.

Solución Calcule del modo siguiente:

$$\begin{aligned} 2\mathbf{v} &= [2(-2), 2(4)] = [-4, 8] \\ \tfrac{1}{2}\mathbf{v} &= [\tfrac{1}{2}(-2), \tfrac{1}{2}(4)] = [-1, 2] \\ -2\mathbf{v} &= [-2(-2), -2(4)] = [4, -8] \end{aligned}$$

Estos vectores se muestran en la figura 1.11.

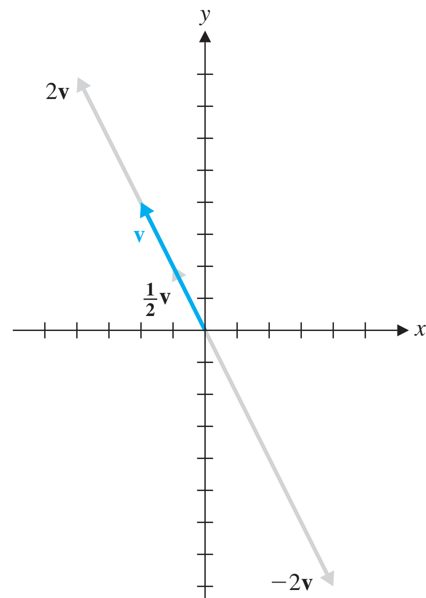


Figura 1.11

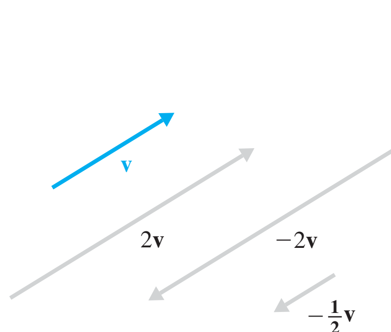


Figura 1.12

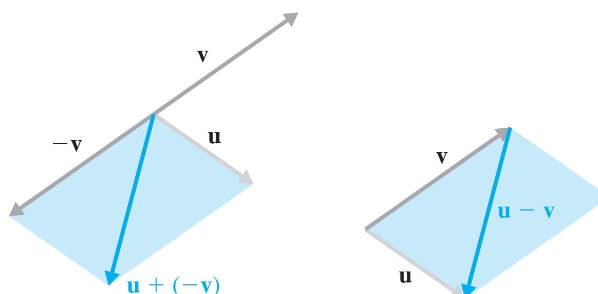


Figura 1.13
Resta de vectores

El término *escalar* proviene de la palabra latina *scala*, que significa "escalera". Los peldaños igualmente espaciados en una escalera sugieren una escala y, en aritmética vectorial, la multiplicación por una constante sólo cambia la escala (o longitud) de un vector. Por ende, las constantes llegaron a conocerse como escalares.

Observe que $c\mathbf{v}$ tiene la misma dirección que $\mathbf{v}$ si $c > 0$ y la dirección opuesta si $c < 0$. También observe que $c\mathbf{v}$ es $|c|$ veces más largo que $\mathbf{v}$. Por esta razón, en el contexto de los vectores, las constantes (esto es: números reales) se conocen como ***escalares***. Como muestra la figura 1.12, cuando se toma en cuenta la traslación de vectores, dos vectores son múltiplos escalares mutuos si y sólo si son ***paralelos***.

Un caso especial de un múltiplo escalar es $(-1)\mathbf{v}$, que se escribe $-\mathbf{v}$ y se llama ***negativo de* v**. Puede usársele para definir la ***resta vectorial***: la ***diferencia*** de **u** y **v** es el vector $\mathbf{u} - \mathbf{v}$ definido por

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$$

La figura 1.13 muestra que $\mathbf{u} - \mathbf{v}$ corresponde a la "otra" diagonal del paralelogramo determinado por **u** y **v**.

Ejemplo 1.4

Si $\mathbf{u} = [1, 2]$ y $\mathbf{v} = [-3, 1]$, entonces $\mathbf{u} - \mathbf{v} = [1 - (-3), 2 - 1] = [4, 1]$.

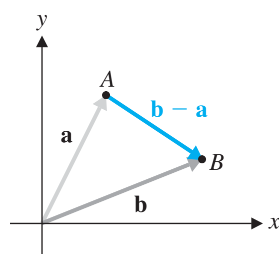


Figura 1.14

La definición de resta en el ejemplo 1.4 también está de acuerdo con la forma en que se calcula un vector como $\overrightarrow{AB}$. Si los puntos A y B corresponden a los vectores **a** y **b** en posición estándar, entonces $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$, como se muestra en la figura 1.14. [Observe que la regla de punta a origen aplicada a este diagrama produce la ecuación $\mathbf{a} + (\mathbf{b} - \mathbf{a}) = \mathbf{b}$. Si accidentalmente dibujara $\mathbf{b} - \mathbf{a}$ con su punta en A en lugar de en B, el diagrama se leería $\mathbf{b} + (\mathbf{b} - \mathbf{a}) = \mathbf{a}$, ¡que claramente está equivocado! Más adelante, en esta sección, se hablará más acerca de las expresiones algebraicas que involucran vectores.]

Vectores en $\mathbb{R}^3$

Todo lo que se ha hecho se extiende con facilidad a tres dimensiones. El conjunto de todas las *ternes ordenadas* de números reales se denota mediante $\mathbb{R}^3$. Puntos y vectores se localizan usando tres ejes coordenados mutuamente perpendiculares que se reúnen en el origen O. Un punto como $A = (1, 2, 3)$ puede localizarse del modo siguiente: primero recorra 1 unidad a lo largo del eje x, luego avance 2 unidades paralelas al eje y y finalmente mueva 3 unidades paralelas al eje z. El vector correspondiente $\mathbf{a} = [1, 2, 3]$ es entonces $\overrightarrow{OA}$, como se muestra en la figura 1.15.

Otra forma de visualizar el vector **a** en $\mathbb{R}^3$ es construir una caja cuyos seis lados estén determinados por los tres planos coordenados (los planos xy, xz y yz) y por tres planos a través del punto $(1, 2, 3)$ paralelos a los planos coordenados. El vector $[1, 2, 3]$ corresponde entonces a la diagonal que va del origen a la esquina opuesta de la caja (vea la figura 1.16).

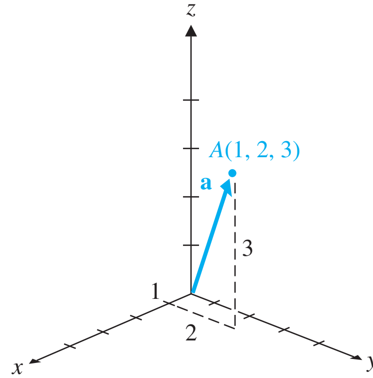


Figura 1.15

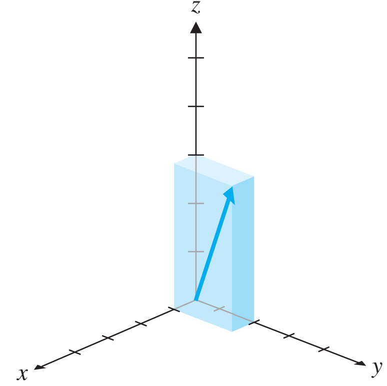


Figura 1.16

Las definiciones "en componentes" de la suma de vectores y la multiplicación escalar se extienden a $\mathbb{R}^3$ en una forma obvia.

Vectores en $\mathbb{R}^n$

En general, $\mathbb{R}^n$ se define como el conjunto de todas las *n-adas ordenadas* de números reales escritos como vectores renglón o columna. Por ende, un vector **v** en $\mathbb{R}^n$ es de la forma

$$[v_1, v_2, \ldots, v_n] \quad \text{o} \quad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

Las entradas individuales de **v** son sus componentes; v_i se llama el componente i-ésimo.

Las definiciones de suma vectorial y multiplicación escalar se extienden a $\mathbb{R}^n$ en la forma obvia: si $\mathbf{u} = [u_1, u_2, \ldots, u_n]$ y $\mathbf{v} = [v_1, v_2, \ldots, v_n]$, el componente i-ésimo de $\mathbf{u} + \mathbf{v}$ es $u_i + v_i$ y el componente i-ésimo de $c\mathbf{v}$ sólo es cv_i.

Dado que en $\mathbb{R}^n$ ya no se pueden dibujar vectores, es importante poder calcularlos. Debe tener cuidado de no suponer que la aritmética vectorial será similar a la aritmética de números reales. Con frecuencia lo es, y los cálculos algebraicos que se hacen con vectores son similares a los que se harían con escalares. Pero, en secciones posteriores se encontrarán situaciones donde el álgebra vectorial es muy *diferente* a la experiencia previa con números reales. Por ello es importante verificar cualquier propiedad algebraica antes de intentar usarla.

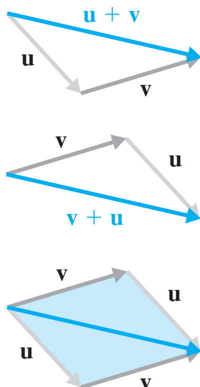


Figura 1.17
$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

Una de tales propiedades es la *conmutatividad* de la adición: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ para los vectores **u** y **v**. Esto ciertamente es verdadero en $\mathbb{R}^2$. Geométricamente, la regla punta a origen muestra que tanto $\mathbf{u} + \mathbf{v}$ como $\mathbf{v} + \mathbf{u}$ son las diagonales principales del paralelogramo determinado por **u** y **v**. (La regla del paralelogramo también refleja esta simetría; vea la figura 1.17.)

Note que la figura 1.17 es simplemente una ilustración de la propiedad $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. No es una demostración, pues no abarca todo caso posible. Por ejemplo, también deben incluirse los casos donde $\mathbf{u} = \mathbf{v}$, $\mathbf{u} = -\mathbf{v}$ y $\mathbf{u} = \mathbf{0}$. (¿Cómo serían los diagramas para estos casos?) Por esta razón, es necesaria una demostración algebraica. Sin embargo, es tan fácil dar una demostración que sea válida en $\mathbb{R}^n$, como dar una que sea válida en $\mathbb{R}^2$.

El siguiente teorema resume las propiedades algebraicas de la suma vectorial y la multiplicación escalar en $\mathbb{R}^n$. Las demostraciones se obtienen de las propiedades correspondientes de los números reales.

Teorema 1.1 **Propiedades algebraicas de los vectores en $\mathbb{R}^n$**

Sean **u**, **v** y **w** vectores en $\mathbb{R}^n$ y sean c y d escalares. Entonces

a. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ — Commutatividad
b. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ — Asociatividad
c. $\mathbf{u} + \mathbf{0} = \mathbf{u}$
d. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
e. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ — Distributividad
f. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ — Distributividad
g. $c(d\mathbf{u}) = (cd)\mathbf{u}$
h. $1\mathbf{u} = \mathbf{u}$

Comentarios

- Las propiedades (c) y (d), junto con la propiedad de conmutatividad (a) implican también que $\mathbf{0} + \mathbf{u} = \mathbf{u}$ y $-\mathbf{u} + \mathbf{u} = \mathbf{0}$.
- Si las propiedades distributivas (e) y (f) se leen de derecha a izquierda, dicen que es posible *factorizar* un escalar común o un vector común a partir de una suma.

La palabra *teorema* se deriva de la palabra griega *theorema*, que a su vez proviene de una palabra que significa "mirar a". Por ende, un teorema se basa en la comprensión que se adquiere cuando se observan los ejemplos y de ellos se extraen propiedades cuya aplicación en general se intenta demostrar. De igual modo, cuando se comprende algo en matemáticas, por ejemplo, la demostración de un teorema, con frecuencia se dice "ya veo".

Demostración Se demuestran las propiedades (a) y (b) y las demostraciones de las propiedades restantes se dejan como ejercicios. Sea $\mathbf{u} = [u_1, u_2, \ldots, u_n]$, $\mathbf{v} = [v_1, v_2, \ldots, v_n]$ y $\mathbf{w} = [w_1, w_2, \ldots, w_n]$.

(a)
$$\begin{aligned}
\mathbf{u} + \mathbf{v} &= [u_1, u_2, \ldots, u_n] + [v_1, v_2, \ldots, v_n] \\
&= [u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n] \\
&= [v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n] \\
&= [v_1, v_2, \ldots, v_n] + [u_1, u_2, \ldots, u_n] \\
&= \mathbf{v} + \mathbf{u}
\end{aligned}$$

La segunda y cuarta igualdades son por la definición de suma vectorial, y la tercera igualdad es por la conmutatividad de la suma de números reales.

(b) La figura 1.18 ilustra la asociatividad en $\mathbb{R}^2$. Algebraicamente, se tiene

$$\begin{aligned}
(\mathbf{u} + \mathbf{v}) + \mathbf{w} &= ([u_1, u_2, \ldots, u_n] + [v_1, v_2, \ldots, v_n]) + [w_1, w_2, \ldots, w_n] \\
&= [u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n] + [w_1, w_2, \ldots, w_n] \\
&= [(u_1 + v_1) + w_1, (u_2 + v_2) + w_2, \ldots, (u_n + v_n) + w_n] \\
&= [u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), \ldots, u_n + (v_n + w_n)] \\
&= [u_1, u_2, \ldots, u_n] + [v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n] \\
&= [u_1, u_2, \ldots, u_n] + ([v_1, v_2, \ldots, v_n] + [w_1, w_2, \ldots, w_n]) \\
&= \mathbf{u} + (\mathbf{v} + \mathbf{w})
\end{aligned}$$

La cuarta igualdad es por la asociatividad de la suma de números reales. Note el uso cuidadoso de los paréntesis.

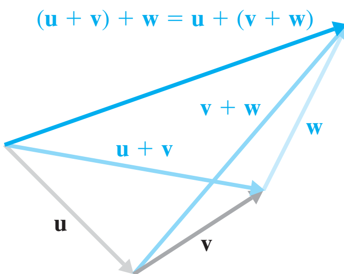


Figura 1.18

Por la propiedad (b) del Teorema 1.1, puede escribirse sin ambigüedades $\mathbf{u} + \mathbf{v} + \mathbf{w}$ sin paréntesis, pues los sumandos pueden agruparse en cualquier forma que se quiera. Por (a), también se pueden reordenar los sumandos si se quiere, por ejemplo, como $\mathbf{w} + \mathbf{u} + \mathbf{v}$. Del mismo modo, las sumas de cuatro o más vectores pueden calcularse sin importar el orden o agrupamiento. En general, si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ son vectores en $\mathbb{R}^n$, tales sumas se escribirán sin paréntesis:

$$\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k$$

El siguiente ejemplo ilustra el uso del Teorema 1.1 en la realización de cálculos algebraicos con vectores.

Ejemplo 1.5

Sean $\mathbf{a}$, $\mathbf{b}$ y $\mathbf{x}$ vectores en $\mathbb{R}^n$.
(a) Simplifique $3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a})$.
(b) Si $5\mathbf{x} - \mathbf{a} = 2(\mathbf{a} + 2\mathbf{x})$, despeje $\mathbf{x}$ en términos de $\mathbf{a}$.

Solución Ambas soluciones se darán con detalle, haciendo con referencia a todas las propiedades en el Teorema 1.1 que se utilicen. Es buena práctica justificar todos los pasos las primeras veces que haga este tipo de cálculos. Sin embargo, una vez se sienta cómodo con las propiedades vectoriales, es aceptable dejar fuera algunos de los pasos intermedios para ahorrar tiempo y espacio.

(a) Comience por insertar paréntesis.

$$\begin{aligned}
3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a}) &= (3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a})) + 2(\mathbf{b} - \mathbf{a}) \\
&= (3\mathbf{a} + (-2\mathbf{a} + 5\mathbf{b})) + (2\mathbf{b} - 2\mathbf{a}) && \text{(a), (e)} \\
&= ((3\mathbf{a} + (-2\mathbf{a})) + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) && \text{(b)} \\
&= ((3 + (-2))\mathbf{a} + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) && \text{(f)} \\
&= (1\mathbf{a} + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) \\
&= ((\mathbf{a} + 5\mathbf{b}) + 2\mathbf{b}) - 2\mathbf{a} && \text{(b), (h)} \\
&= (\mathbf{a} + (5\mathbf{b} + 2\mathbf{b})) - 2\mathbf{a} && \text{(b)} \\
&= (\mathbf{a} + (5 + 2)\mathbf{b}) - 2\mathbf{a} && \text{(f)} \\
&= (7\mathbf{b} + \mathbf{a}) - 2\mathbf{a} && \text{(a)} \\
&= 7\mathbf{b} + (\mathbf{a} - 2\mathbf{a}) && \text{(b)} \\
&= 7\mathbf{b} + (1 - 2)\mathbf{a} && \text{(f), (h)} \\
&= 7\mathbf{b} + (-1)\mathbf{a} \\
&= 7\mathbf{b} - \mathbf{a}
\end{aligned}$$

¡Puede ver por qué estará de acuerdo en omitir algunos de estos pasos! En la práctica es aceptable simplificar esta secuencia de pasos como

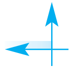
$$\begin{aligned}
3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a}) &= 3\mathbf{a} + 5\mathbf{b} - 2\mathbf{a} + 2\mathbf{b} - 2\mathbf{a} \\
&= (3\mathbf{a} - 2\mathbf{a} - 2\mathbf{a}) + (5\mathbf{b} + 2\mathbf{b}) \\
&= -\mathbf{a} + 7\mathbf{b}
\end{aligned}$$

o incluso hacer la mayoría de los cálculos mentalmente.

(b) En detalle, se tiene

$$\begin{aligned}
5\mathbf{x} - \mathbf{a} &= 2(\mathbf{a} + 2\mathbf{x}) & \\
5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + 2(2\mathbf{x}) & \text{(e)}\\
5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + (2 \cdot 2)\mathbf{x} & \text{(g)}\\
5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + 4\mathbf{x} & \\
(5\mathbf{x} - \mathbf{a}) - 4\mathbf{x} &= (2\mathbf{a} + 4\mathbf{x}) - 4\mathbf{x} & \\
(-\mathbf{a} + 5\mathbf{x}) - 4\mathbf{x} &= 2\mathbf{a} + (4\mathbf{x} - 4\mathbf{x}) & \text{(a), (b)}\\
-\mathbf{a} + (5\mathbf{x} - 4\mathbf{x}) &= 2\mathbf{a} + \mathbf{0} & \text{(b), (d)}\\
-\mathbf{a} + (5 - 4)\mathbf{x} &= 2\mathbf{a} & \text{(f), (c)}\\
-\mathbf{a} + (1)\mathbf{x} &= 2\mathbf{a} & \\
\mathbf{a} + (-\mathbf{a} + \mathbf{x}) &= \mathbf{a} + 2\mathbf{a} & \text{(h)}\\
(\mathbf{a} + (-\mathbf{a})) + \mathbf{x} &= (1 + 2)\mathbf{a} & \text{(b), (f)}\\
\mathbf{0} + \mathbf{x} &= 3\mathbf{a} & \text{(d)}\\
\mathbf{x} &= 3\mathbf{a} & \text{(c)}
\end{aligned}$$

De nuevo, usualmente se omitirán la mayoría de estos pasos.



Combinaciones lineales y coordenadas

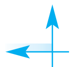
Se dice que un vector que sea una suma de múltiplos escalares de otros vectores es una *combinación lineal* de dichos vectores. A continuación se presenta la definición formal.

Definición Un vector $\mathbf{v}$ es una ***combinación lineal*** de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ si existen escalares $c_1, c_2, \ldots, c_k$ tales que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$. Los escalares $c_1, c_2, \ldots, c_k$ se llaman ***coeficientes*** de la combinación lineal.

Ejemplo 1.6

El vector $\begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$ es una combinación lineal de $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 5 \\ -4 \\ 0 \end{bmatrix}$, pues

$$3\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} - \begin{bmatrix} 5 \\ -4 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$$



Comentario Determinar si un vector dado es una combinación lineal de otros vectores es un problema que se abordará en el capítulo 2.

En $\mathbb{R}^2$, es posible bosquejar combinaciones lineales de dos vectores (no paralelos) de manera muy conveniente.

Ejemplo 1.7

Sean $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Puede usar $\mathbf{u}$ y $\mathbf{v}$ para ubicar un nuevo conjunto de ejes (en la misma forma que $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ ubican los ejes coordenados estándar). Puede usar

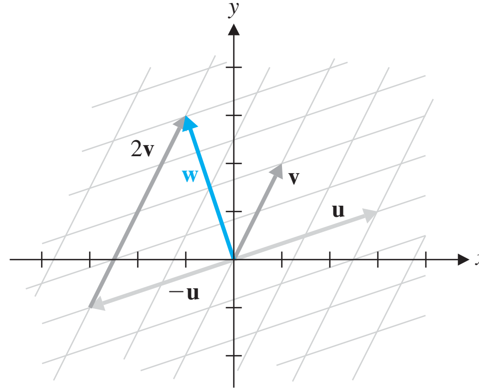


Figura 1.19

estos nuevos ejes para determinar una ***cuadrícula coordenada*** que permitirá ubicar fácilmente combinaciones lineales de **u** y **v**.

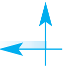
Como muestra la figura 1.19, **w** puede ubicarse comenzando en el origen y avanzar –**u** seguido por 2**v**. Esto es,

$$\mathbf{w} = -\mathbf{u} + 2\mathbf{v}$$

Se dice que las coordenadas de **w** con respecto a **u** y **v** son -1 y 2. (Note que esta es sólo otra forma de pensar en los coeficientes de la combinación lineal.) Se concluye que

$$\mathbf{w} = -\begin{bmatrix} 3 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

(Observe que -1 y 3 son las coordenadas de **w** con respecto a $\mathbf{e}_1$ y $\mathbf{e}_2$.)



Cambiar de los ejes coordenados estándar a los alternativos es una idea útil. Tiene aplicaciones en química y geología, pues las estructuras moleculares y cristalinas con frecuencia no se adaptan a una cuadrícula rectangular. Es una idea que se encontrará repetidamente en este libro.

Vectores binarios y aritmética modular

También encontrará un tipo de vector que no tiene interpretación geométrica, al menos no con el uso de la geometría euclidiana. Las computadoras representan los datos en términos de 0 y 1 (que se puede interpretar como apagado/encendido, cerrado/abierto, falso/verdadero o no/sí). Los ***vectores binarios*** son vectores cuyos componentes son un 0 o un 1. Como verá en la sección 1.4, tales vectores surgen naturalmente en el estudio de muchos tipos de códigos.

En este escenario deben modificarse las reglas usuales de la aritmética, pues el resultado de cada cálculo que involucra escalares debe ser un 0 o un 1. A continuación se proporcionan las reglas modificadas para suma y multiplicación.

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

La única curiosidad aquí es la regla de que $1 + 1 = 0$. Esto no es tan extraño como parece; si sustituye 0 con la palabra “par” y 1 con la palabra “impar”, estas tablas simplemente

resumen las familiares *reglas de paridad* para la suma y multiplicación de enteros pares e impares. Por ejemplo, $1 + 1 = 0$ expresa el hecho de que la suma de dos enteros impares es un entero par. Con estas reglas, el conjunto de escalares $\{0, 1\}$ se denota mediante $\mathbb{Z}_2$ y se llama el conjunto de ***enteros módulo*** **2**.

Ejemplo 1.8

En; $\mathbb{Z}_2$, $1 + 1 + 0 + 1 = 1$ y $1 + 1 + 1 + 1 = 0$. (Estos cálculos ilustran nuevamente las reglas de paridad: la suma de tres impares y un par es impar; la suma de cuatro impares es par.)

Se usa el término *longitud* de manera diferente a la forma en como se usa en $\mathbb{R}^n$. Esto no debe confundirlo, pues no hay noción *geométrica* de longitud para vectores binarios.

Con $\mathbb{Z}_2$ como el conjunto de escalares, ahora las reglas anteriores se extienden a los vectores. El conjunto de todas las n-adas de 0 y 1 (con toda la aritmética realizada en módulo 2) se denota mediante $\mathbb{Z}_2^n$. Los vectores en $\mathbb{Z}_2^n$ se llaman ***vectores binarios de longitud n***.

Ejemplo 1.9

Los vectores en $\mathbb{Z}_2^2$ son $[0, 0]$, $[0, 1]$, $[1, 0]$ y $[1, 1]$. (En general, ¿cuántos vectores contiene $\mathbb{Z}_2^n$?)

Ejemplo 1.10

Sean; $\mathbf{u} = [1, 1, 0, 1, 0]$ y $\mathbf{v} = [0, 1, 1, 1, 0]$ dos vectores binarios de longitud 5. Encuentre $\mathbf{u} \cdot \mathbf{v}$.

Solución El cálculo de $\mathbf{u} \cdot \mathbf{v}$ tiene lugar en $\mathbb{Z}_2$, de modo que se tiene

$$\begin{aligned}\mathbf{u} \cdot \mathbf{v} &= 1 \cdot 0 + 1 \cdot 1 + 0 \cdot 1 + 1 \cdot 1 + 0 \cdot 0 \\ &= 0 + 1 + 0 + 1 + 0 \\ &= 0\end{aligned}$$

Es posible generalizar lo que se acaba de hacer para vectores binarios a vectores cuyos componentes se tomen de un conjunto finito $\{0, 1, 2, \ldots, k\}$ para $k \geq 2$. Para hacerlo, primero debe ampliarse la idea de aritmética binaria.

Ejemplo 1.11

Los ***enteros módulo*** **3** es el conjunto $\mathbb{Z}_3 = \{0, 1, 2\}$ con suma y multiplicación dadas por las siguientes tablas:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Observe que el resultado de cada suma y multiplicación pertenece al conjunto $\{0, 1, 2\}$; se dice que $\mathbb{Z}$ es ***cerrado*** con respecto a las operaciones de suma y multiplicación. Tal vez sea más sencillo pensar en este conjunto en términos de un reloj de 3 horas con 0, 1 y 2 en su carátula, como se muestra en la figura 1.20.

El cálculo $1 + 2 = 0$ se traduce del modo siguiente: 2 horas después de la 1 en punto es 0 en punto. Así como 24:00 y 12:00 son lo mismo en un reloj de 12 horas, del mismo modo 3 y 0 son equivalentes en este reloj de 3 horas. Igualmente, todos los múltiplos de 3, positivos y negativos, aquí son equivalentes a 0; 1 es equivalente a cualquier número que sea 1 más que un múltiplo de 3 (como -2, 4 y 7); y 2 es equivalente a cualquier nú-

mero que sea 2 más que un múltiplo de 3 (como –1, 5 y 8). Se puede visualizar la recta numérica como enredada en torno a un círculo, como se muestra en la figura 1.21.

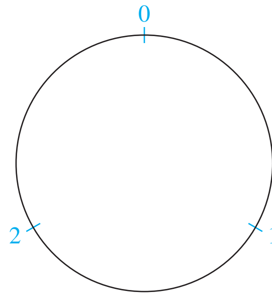


Figura 1.20
Módulo aritmético

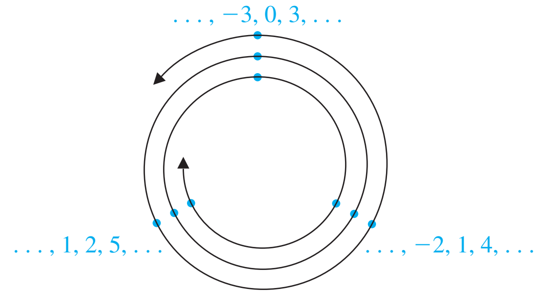


Figura 1.21

Ejemplo 1.12

¿A qué es equivalente 3548 en $\mathbb{Z}_3$?

Solución Esto es lo mismo que preguntar dónde se encuentra 3548 en el reloj de 3 horas. La clave es calcular cuán lejos está el número del múltiplo de 3 más cercano (más pequeño); es decir, es necesario conocer el *residuo* cuando 3548 se divide entre 3. Por división larga, se encuentra que $3548 = 3 \cdot 1182 + 2$, de modo que el residuo es 2. Por tanto, 3548 es equivalente a 2 en $\mathbb{Z}_3$.

En cursos de álgebra abstracta y teoría de números, que exploran este concepto con mayor detalle, la equivalencia anterior con frecuencia se escribe como $3548 = 2 \pmod 3$ o $3548 \equiv 2 \pmod 3$, donde $\equiv$ se lee "es congruente con". Aquí no se usará esta notación o terminología.

Ejemplo 1.13

En $\mathbb{Z}_3$, calcule $2 + 2 + 1 + 2$.

Solución 1 Use las mismas ideas que en el ejemplo 1.12. La suma ordinaria es $2 + 2 + 1 + 2 = 7$, que es 1 más que 6, de modo que la división entre 3 deja un residuo de 1. Por ende, $2 + 2 + 1 + 2 = 1$ en $\mathbb{Z}_3$.

Solución 2 Una mejor forma de realizar este cálculo es hacerlo paso a paso completamente en $\mathbb{Z}_3$.

$$\begin{aligned} 2 + 2 + 1 + 2 &= (2 + 2) + 1 + 2 \\ &= 1 + 1 + 2 \\ &= (1 + 1) + 2 \\ &= 2 + 2 \\ &= 1 \end{aligned}$$

Aquí se usaron paréntesis para agrupar los términos elegidos para combinar. Podría acelerar las cosas al combinar simultáneamente los primeros dos y los últimos dos términos:

$$\begin{aligned} (2 + 2) + (1 + 2) &= 1 + 0 \\ &= 1 \end{aligned}$$

La multiplicación repetida se puede manejar de manera similar. La idea es usar las tablas de suma y multiplicación para reducir el resultado de cada cálculo a 0, 1 o 2.



La extensión de estas ideas a los vectores es directa.

Ejemplo 1.14

En $\mathbb{Z}_3^5$, sean $\mathbf{u} = [2, 2, 0, 1, 2]$ y $\mathbf{v} = [1, 2, 2, 2, 1]$. Entonces

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= 2\cdot 1 + 2\cdot 2 + 0\cdot 2 + 1\cdot 2 + 2\cdot 1\\ &= 2 + 1 + 0 + 2 + 2\\ &= 1\end{aligned}$$

Los vectores en $\mathbb{Z}_3^5$, se conocen como ***vectores ternarios de longitud* 5.**

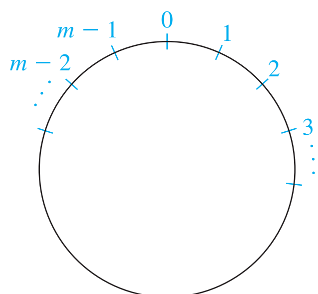


Figura 1.22
Módulo aritmético m

En general, se tiene el conjunto $\mathbb{Z}_m = \{0, 1, 2, \ldots, m - 1\}$ de ***enteros módulo m*** (que corresponde a un reloj de m horas, como se muestra en la figura 1.22). Un vector de longitud n cuyas entradas están en $\mathbb{Z}_m$ se llama ***vector m-ario de longitud n***. El conjunto de todos los vectores m-arios de longitud n se denota por $\mathbb{Z}_m^n$.

Ejercicios 1.1

1. Dibuje los siguientes vectores en posición estándar en $\mathbb{R}^2$.

(a) $\mathbf{a} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ **(b)** $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

(c) $\mathbf{c} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ **(d)** $\mathbf{d} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

2. Dibuje los vectores del ejercicio 1 con sus orígenes en el punto $(1, -3)$.

3. Dibuje los siguientes vectores en posición estándar en $\mathbb{R}^3$.

(a) $\mathbf{a} = [0, 2, 0]$ **(b)** $\mathbf{b} = [3, 2, 1]$
(c) $\mathbf{c} = [1, -2, 1]$ **(d)** $\mathbf{d} = [-1, -1, -2]$

4. Si los vectores en el ejercicio 3 se trasladan de modo que sus puntas estén en el punto $(1, 2, 3)$, encuentre los puntos que correspondan a sus orígenes.

5. Para cada uno de los siguientes pares de puntos, dibuje el vector $\overrightarrow{AB}$. Luego calcule y vuelva a dibujar $\overrightarrow{AB}$ como un vector en posición estándar.

(a) $A = (1, -1), B = (4, 2)$
(b) $A = (0, -2), B = (2, -1)$
(c) $A = (2, \frac{3}{2}), B = (\frac{1}{2}, 3)$
(d) $A = (\frac{1}{3}, \frac{1}{3}), B = (\frac{1}{6}, \frac{1}{2})$

6. Un paseante camina 4 km al norte y luego 5 km al noreste. Dibuje los vectores desplazamiento que representen el viaje del caminante y dibuje un vector que represente el desplazamiento neto del caminante desde el punto de partida.

Los ejercicios 7-10 se refieren a los vectores en el ejercicio 1. Calcule los vectores indicados y también muestre cómo pueden obtenerse geométricamente los resultados.

7. $\mathbf{a} + \mathbf{b}$ **8.** $\mathbf{b} + \mathbf{c}$

9. $\mathbf{d} - \mathbf{c}$ **10.** $\mathbf{a} - \mathbf{d}$

Los ejercicios 11 y 12 se refieren a los vectores en el ejercicio 3. Calcule los vectores indicados.

11. $2\mathbf{a} + 3\mathbf{c}$ **12.** $2\mathbf{c} - 3\mathbf{b} - \mathbf{d}$

13. Encuentre los componentes de los vectores $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$, donde $\mathbf{u}$ y $\mathbf{v}$ son como se muestra en la figura 1.23.

14. En la figura 1.24, A, B, C, D, E y F son los vértices de un hexágono regular con centro en el origen.

Exprese cada uno de los siguientes vectores en términos de $\mathbf{a} = \overrightarrow{OA}$ y $\mathbf{b} = \overrightarrow{OB}$:

(a) $\overrightarrow{AB}$ **(b)** $\overrightarrow{BC}$

(c) $\overrightarrow{AD}$ **(d)** $\overrightarrow{CF}$

(e) $\overrightarrow{AC}$ **(f)** $\overrightarrow{BC} + \overrightarrow{DE} + \overrightarrow{FA}$

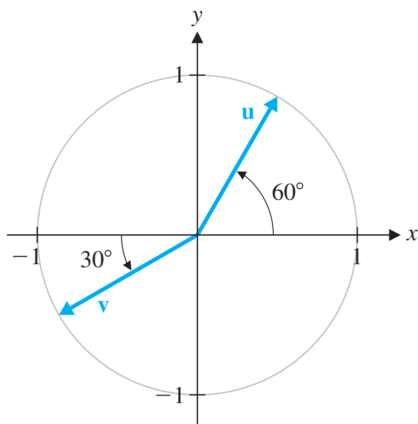


Figura 1.23

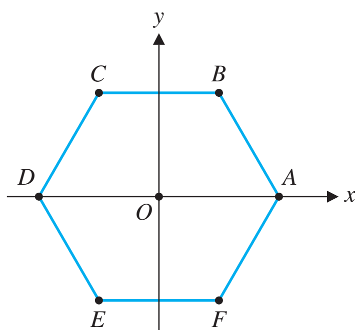


Figura 1.24

En los ejercicios 15 y 16, simplifique la expresión vectorial dada. Indique cuáles propiedades usó del Teorema 1.1.

15. $2(\mathbf{a} - 3\mathbf{b}) + 3(2\mathbf{b} + \mathbf{a})$

16. $-3(\mathbf{a} - \mathbf{c}) + 2(\mathbf{a} + 2\mathbf{b}) + 3(\mathbf{c} - \mathbf{b})$

En los ejercicios 17 y 18, despeje el vector **x** *en términos de los vectores* **a** *y* **b**.

17. $\mathbf{x} - \mathbf{a} = 2(\mathbf{x} - 2\mathbf{a})$

18. $\mathbf{x} + 2\mathbf{a} - \mathbf{b} = 3(\mathbf{x} + \mathbf{a}) - 2(2\mathbf{a} - \mathbf{b})$

En los ejercicios 19 y 20, dibuje los ejes coordenados con respecto a **u** *y* **v** *y localice* **w**.

19. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w} = 2\mathbf{u} + 3\mathbf{v}$

20. $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{w} = -\mathbf{u} - 2\mathbf{v}$

En los ejercicios 21 y 22, dibuje los ejes coordenados estándar en el mismo diagrama que los ejes con respecto a **u** *y* **v**. *Úselos para encontrar* **w** *como una combinación lineal de* **u** *y* **v**.

21. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$

22. $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 2 \\ 9 \end{bmatrix}$

23. Dibuje diagramas para ilustrar las propiedades (d) y (e) del Teorema 1.1.

24. Proporcione demostraciones algebraicas de las propiedades (d) a (g) del Teorema 1.1.

En los ejercicios 25-28, **u** *y* **v** *son vectores binarios. Encuentre* $\mathbf{u} + \mathbf{v}$ *y* $\mathbf{u} \cdot \mathbf{v}$ *en cada caso.*

25. $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ **26.** $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

27. $\mathbf{u} = [1, 0, 1, 1], \mathbf{v} = [1, 1, 1, 1]$

28. $\mathbf{u} = [1, 1, 0, 1, 0], \mathbf{v} = [0, 1, 1, 1, 0]$

29. Escriba las tablas de suma y multiplicación para $\mathbb{Z}_4$.

30. Escriba las tablas de suma y multiplicación para $\mathbb{Z}_5$.

En los ejercicios 31-43, realice los cálculos indicados.

31. $2 + 2 + 2$ en $\mathbb{Z}_3$ **32.** $2 \cdot 2 \cdot 2$ en $\mathbb{Z}_3$

33. $2(2 + 1 + 2)$ en $\mathbb{Z}_3$ **34.** $3 + 1 + 2 + 3$ en $\mathbb{Z}_4$

35. $2 \cdot 3 \cdot 2$ en $\mathbb{Z}_4$ **36.** $3(3 + 3 + 2)$ en $\mathbb{Z}_4$

37. $2 + 1 + 2 + 2 + 1$ en $\mathbb{Z}_3$, $\mathbb{Z}_4$, y $\mathbb{Z}_5$

38. $(3 + 4)(3 + 2 + 4 + 2)$ en $\mathbb{Z}_5$

39. $8(6 + 4 + 3)$ en $\mathbb{Z}_9$ **40.** 2^{100} en $\mathbb{Z}_{11}$

41. $[2, 1, 2] + [2, 0, 1]$ en $\mathbb{Z}_3^3$ **42.** $[2, 1, 2] \cdot [2, 2, 1]$ en $\mathbb{Z}_3^3$

43. $[2, 0, 3, 2] \cdot ([3, 1, 1, 2] + [3, 3, 2, 1])$ en $\mathbb{Z}_4^4$ y en $\mathbb{Z}_5^4$

En los ejercicios 44-55, resuelva la ecuación dada o indique que no hay solución.

44. $x + 3 = 2$ en $\mathbb{Z}_5$ **45.** $x + 5 = 1$ en $\mathbb{Z}_6$

46. $2x = 1$ en $\mathbb{Z}_3$ **47.** $2x = 1$ en $\mathbb{Z}_4$

48. $2x = 1$ en $\mathbb{Z}_5$ **49.** $3x = 4$ en $\mathbb{Z}_5$

50. $3x = 4$ en $\mathbb{Z}_6$ **51.** $6x = 5$ en $\mathbb{Z}_8$

52. $8x = 9$ en $\mathbb{Z}_{11}$ **53.** $2x + 3 = 2$ en $\mathbb{Z}_5$

54. $4x + 5 = 2$ en $\mathbb{Z}_6$ **55.** $6x + 3 = 1$ en $\mathbb{Z}_8$

56. **(a)** ¿Para qué valores de a, $x + a = 0$ tiene solución en $\mathbb{Z}_5$?
(b) ¿Para qué valores de a y b, $x + a = b$ tiene solución en $\mathbb{Z}_6$?
(c) ¿Para qué valores de a, b y m, $x + a = b$ tiene solución en $\mathbb{Z}_m$?

57. **(a)** ¿Para qué valores de a, $ax = 1$ tiene solución en $\mathbb{Z}_5$?
(b) ¿Para qué valores de a, $ax = 1$ tiene solución en $\mathbb{Z}_6$?
(c) ¿Para qué valores de a y m, $ax = 1$ tiene solución en $\mathbb{Z}_m$?

1.2 Longitud y ángulo: el producto punto

Es muy sencillo reformular los conceptos geométricos familiares de longitud, distancia y ángulo en términos de vectores. Hacerlo permitirá usar estas importantes y poderosas ideas en escenarios más generales que $\mathbb{R}^2$ y $\mathbb{R}^3$. En capítulos posteriores, estas simples herramientas geométricas se usarán para resolver una gran variedad de problemas que surgen en aplicaciones, ¡aun cuando aparentemente no haya ninguna geometría!

El producto punto

Las versiones vectoriales de longitud, distancia y ángulo pueden describirse usando la noción de producto punto de dos vectores.

Definición Si

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

entonces el ***producto punto*** $\mathbf{u} \cdot \mathbf{v}$ de $\mathbf{u}$ y $\mathbf{v}$ se define mediante

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

En palabras, $\mathbf{u} \cdot \mathbf{v}$ es la suma de los productos de los componentes correspondientes de $\mathbf{u}$ y $\mathbf{v}$. Es importante notar un par de cosas acerca de este “producto” que se acaba de definir: primero, $\mathbf{u}$ y $\mathbf{v}$ deben tener el mismo número de componentes. Segundo, el producto punto $\mathbf{u} \cdot \mathbf{v}$ es un *número*, no otro vector. (Es por esto que $\mathbf{u} \cdot \mathbf{v}$ en ocasiones se conoce como el ***producto escalar*** de $\mathbf{u}$ y $\mathbf{v}$.) El producto punto de vectores en $\mathbb{R}^n$ es un caso especial e importante de la noción más general de ***producto interno***, que se explorará en el capítulo 7.

Ejemplo 1.15

Calcule $\mathbf{u} \cdot \mathbf{v}$ cuando $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \\ 2 \end{bmatrix}$.

Solución $\mathbf{u} \cdot \mathbf{v} = 1 \cdot (-3) + 2 \cdot 5 + (-3) \cdot 2 = 1$

Note que si en el ejemplo 1.15 se hubiera calculado $\mathbf{v} \cdot \mathbf{u}$, se habría obtenido

$$\mathbf{v} \cdot \mathbf{u} = (-3) \cdot 1 + 5 \cdot 2 + 2 \cdot (-3) = 1$$

Es algo que, en general, $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$, pues los productos individuales de los componentes se conmutan. Esta propiedad de conmutatividad es una de las propiedades del producto punto que se usarán repetidamente. Las principales propiedades del producto punto se resumen en el Teorema 1.2.

Teorema 1.2

Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores en $\mathbb{R}^n$, y sea c un escalar. Entonces

a. $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$ — Conmutatividad
b. $\mathbf{u}\cdot(\mathbf{v}+\mathbf{w}) = \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}$ — Distributividad
c. $(c\mathbf{u})\cdot\mathbf{v} = c(\mathbf{u}\cdot\mathbf{v})$
d. $\mathbf{u}\cdot\mathbf{u} \geq 0$ y $\mathbf{u}\cdot\mathbf{u} = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$

Demostración Se demostrarán (a) y (c), y la demostración de las propiedades restantes se dejará para los ejercicios.

(a) Al aplicar la definición de producto punto a $\mathbf{u}\cdot\mathbf{v}$ y $\mathbf{v}\cdot\mathbf{u}$ se obtiene

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n\\ &= v_1u_1 + v_2u_2 + \cdots + v_nu_n\\ &= \mathbf{v}\cdot\mathbf{u}\end{aligned}$$

donde la segunda igualdad es consecuencia del hecho de que la multiplicación de números reales es conmutativa.

(c) Con las definiciones de multiplicación escalar y producto punto, se tiene

$$\begin{aligned}(c\mathbf{u})\cdot\mathbf{v} &= [cu_1, cu_2, \ldots, cu_n]\cdot[v_1, v_2, \ldots, v_n]\\ &= cu_1v_1 + cu_2v_2 + \cdots + cu_nv_n\\ &= c(u_1v_1 + u_2v_2 + \cdots + u_nv_n)\\ &= c(\mathbf{u}\cdot\mathbf{v})\end{aligned}$$

Comentarios

- La propiedad (b) puede leerse de derecha a izquierda, en cuyo caso se dice que se puede factorizar un vector común $\mathbf{u}$ a partir de una suma de productos punto. Esta propiedad también tienen un análogo de "mano derecha" que procede de las propiedades (b) y (a) juntas: $(\mathbf{v}+\mathbf{w})\cdot\mathbf{u} = \mathbf{v}\cdot\mathbf{u} + \mathbf{w}\cdot\mathbf{u}$.
- La propiedad (c) puede extenderse para dar $\mathbf{u}\cdot(c\mathbf{v}) = c(\mathbf{u}\cdot\mathbf{v})$ (ejercicio 58). Esta versión extendida de (c) dice en esencia que, al tomar un múltiplo escalar de un producto punto de vectores, el escalar puede combinarse primero con cualquier vector que sea más conveniente. Por ejemplo,

$$\left(\tfrac{1}{2}[-1,-3,2]\right)\cdot[6,-4,0] = [-1,-3,2]\cdot\left(\tfrac{1}{2}[6,-4,0]\right) = [-1,-3,2]\cdot[3,-2,0] = 3$$

Con este enfoque se evita la introducción de fracciones en los vectores, como ocurriría en el agrupamiento original.

- La segunda parte de (d) usa el conectivo lógico *si y sólo si*. El Apéndice A discute esta frase con más detalle, pero por el momento note solamente que la frase señala una *doble implicación*, a saber

si $\mathbf{u} = \mathbf{0}$, entonces $\mathbf{u}\cdot\mathbf{u} = 0$

y

si $\mathbf{u}\cdot\mathbf{u} = 0$, entonces $\mathbf{u} = \mathbf{0}$

El Teorema 1.2 muestra que algunos aspectos del álgebra de vectores recuerdan el álgebra de números. El siguiente ejemplo muestra que en ocasiones es posible encontrar análogos vectoriales de identidades familiares.

Ejemplo 1.16 Demuestre que $(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \mathbf{u}\cdot\mathbf{u} + 2(\mathbf{u}\cdot\mathbf{v}) + \mathbf{v}\cdot\mathbf{v}$ para todos los vectores **u** y **v** en $\mathbb{R}^n$.

Solución

$$\begin{aligned}(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) &= (\mathbf{u}+\mathbf{v})\cdot\mathbf{u} + (\mathbf{u}+\mathbf{v})\cdot\mathbf{v}\\ &= \mathbf{u}\cdot\mathbf{u} + \mathbf{v}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{v}\\ &= \mathbf{u}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{v}\\ &= \mathbf{u}\cdot\mathbf{u} + 2(\mathbf{u}\cdot\mathbf{v}) + \mathbf{v}\cdot\mathbf{v}\end{aligned}$$

(Identifique las partes del Teorema 1.2 que se usaron en cada paso.)

Longitud

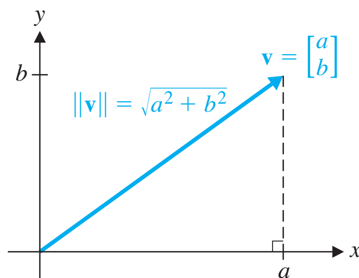


Figura 1.25

Para ver cómo el producto punto desempeña un papel en el cálculo de longitudes, recuerde cómo se calculan longitudes en el plano. El teorema de Pitágoras es todo lo que necesita.

En $\mathbb{R}^2$, la longitud del vector $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ es la distancia desde el origen hasta el punto (a, b) que, por el teorema de Pitágoras, está dada por $\sqrt{a^2 + b^2}$, como en la figura 1.25. Observe que $a^2 + b^2 = \mathbf{v}\cdot\mathbf{v}$. Esto conduce a la siguiente definición.

Definición La ***longitud*** (o ***norma***) de un vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ en $\mathbb{R}^n$ es el escalar no negativo $\|\mathbf{v}\|$ definido por

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

En palabras, la longitud de un vector es la raíz cuadrada de la suma de los cuadrados de sus componentes. Note que la raíz cuadrada de $\mathbf{v}\cdot\mathbf{v}$ siempre está definida, pues $\mathbf{v}\cdot\mathbf{v} \geq 0$ por el Teorema 1.2(d). Note también que la definición puede reescribirse para dar $\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$, que será útil para probar más propiedades del producto punto y las longitudes de vectores.

Ejemplo 1.17

$$\|[2, 3]\| = \sqrt{2^2 + 3^2} = \sqrt{13}$$

El Teorema 1.3 menciona algunas de las principales propiedades de la longitud vectorial.

Teorema 1.3 Sea **v** un vector en $\mathbb{R}^n$ y sea c un escalar. Entonces

a. $\|\mathbf{v}\| = 0$ si y sólo si $\mathbf{v} = \mathbf{0}$
b. $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$

Demostración La propiedad (a) es consecuencia inmediata del Teorema 1.2(d). Para demostrar (b), se tiene

$$\|c\mathbf{v}\|^2 = (c\mathbf{v})\cdot(c\mathbf{v}) = c^2(\mathbf{v}\cdot\mathbf{v}) = c^2\|\mathbf{v}\|^2$$

usando el Teorema 1.2(c). Al sacar la raíz cuadrada de ambos lados, y con el hecho de que $\sqrt{c^2} = |c|$ para cualquier número real c, se obtiene el resultado.

Un vector de longitud 1 se llama ***vector unitario***. En $\mathbb{R}^2$, el conjunto de todos los vectores unitarios puede identificarse con el *círculo unitario*, el círculo de radio 1 con centro en el origen (vea la figura 1.26). Dado cualquier vector $\mathbf{v}$ distinto de cero, siempre es posible encontrar un vector unitario en la misma dirección que $\mathbf{v}$ al dividir $\mathbf{v}$ por su propia longitud (o, de manera equivalente, al *multiplicar* por $1/\|\mathbf{v}\|$). Esto se puede demostrar algebraicamente al usar la propiedad (b) del Teorema 1.3 anterior: si $\mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v}$, entonces

$$\|\mathbf{u}\| = \|(1/\|\mathbf{v}\|)\mathbf{v}\| = |1/\|\mathbf{v}\|| \, \|\mathbf{v}\| = (1/\|\mathbf{v}\|)\|\mathbf{v}\| = 1$$

y $\mathbf{u}$ está en la misma dirección que $\mathbf{v}$, ya que $1/\|\mathbf{v}\|$ es un escalar positivo. Encontrar un vector unitario en la misma dirección con frecuencia se conoce como ***normalizar*** un vector (vea la figura 1.27).

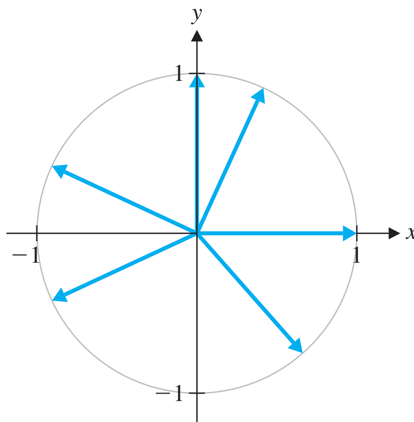


Figura 1.26
Vectores unitarios en $\mathbb{R}^2$

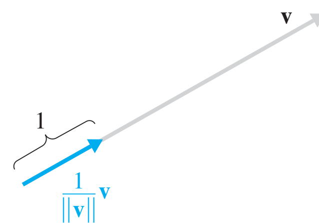


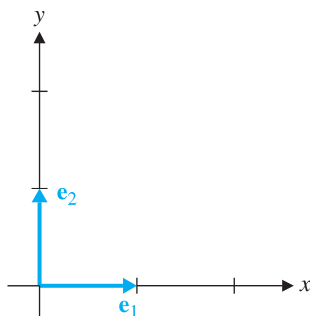
Figura 1.27
Normalización de un vector

Ejemplo 1.18

En $\mathbb{R}^2$, sea $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Entonces $\mathbf{e}_1$ y $\mathbf{e}_2$ son vectores unitarios, ya que la suma de los cuadrados de sus componentes es 1 en cada caso. De igual modo, en $\mathbb{R}^3$ pueden construirse vectores unitarios

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Observe en la figura 1.28 que estos vectores sirven para localizar los ejes coordenados positivos en $\mathbb{R}^2$ y $\mathbb{R}^3$.



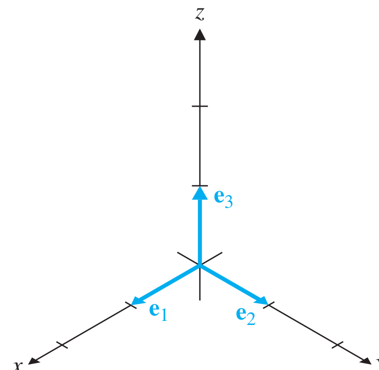


Figura 1.28
Vectores unitarios estándar en $\mathbb{R}^2$ y $\mathbb{R}^3$

En general, en $\mathbb{R}^n$, se definen vectores unitarios $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, donde $\mathbf{e}_i$ tiene 1 en su i-ésimo componente y ceros en cualquier otra parte. Tales vectores surgen de manera repetida en álgebra lineal y se llaman ***vectores unitarios estándar***.

Ejemplo 1.19

Normalice el vector $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

Solución $\|\mathbf{v}\| = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{14}$, de modo que un vector unitario en la misma dirección que $\mathbf{v}$ está dado por

$$\mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v} = (1/\sqrt{14})\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{14} \\ -1/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$$

Puesto que la propiedad (b) del Teorema 1.3 describe cómo se comporta la longitud con respecto a la multiplicación escalar, la curiosidad natural sugiere que se pregunte si la longitud y la suma vectorial son compatibles. Sería excelente tener una identidad como $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$, pero para casi cualquier elección de vectores $\mathbf{u}$ y $\mathbf{v}$ esto resulta ser falso. [Vea el ejercicio 52(a).] Sin embargo, no todo está perdido, pues resulta que, si se sustituye el signo $=$ con $\leq$, la desigualdad resultante es verdadera. La demostración de este famoso e importante resultado, la desigualdad del triángulo, se apoya en otra importante desigualdad, la de Cauchy-Schwarz, que se probará y discutirá con más detalle en el capítulo 7.

Teorema 1.4

La desigualdad de Cauchy-Schwarz

Para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$,

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$$

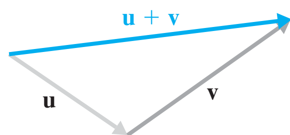


Figura 1.29
La desigualdad del triángulo

Vea en los ejercicios 71 y 72 los planteamientos algebraico y geométrico para la demostración de esta desigualdad.

En $\mathbb{R}^2$ o $\mathbb{R}^3$, donde se puede usar geometría, es claro a partir de un diagrama como el de la figura 1.29 que $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ para todos los vectores $\mathbf{u}$ y $\mathbf{v}$. Ahora se demostrará que esto es verdadero en general.

Teorema 1.5

La desigualdad del triángulo

Para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$



Demostración Dado que ambos lados de la desigualdad son no negativos, demostrar que el *cuadrado* del lado izquierdo es menor o igual que el *cuadrado* del lado derecho equivale a demostrar el teorema. (¿Por qué?) Calcule

$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) \\
&= \mathbf{u}\cdot\mathbf{u} + 2(\mathbf{u}\cdot\mathbf{v}) + \mathbf{v}\cdot\mathbf{v} && \text{Por el ejemplo 1.9} \\
&\le \|\mathbf{u}\|^2 + 2|\mathbf{u}\cdot\mathbf{v}| + \|\mathbf{v}\|^2 \\
&\le \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 && \text{Por Cauchy-Schwarz} \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2
\end{aligned}$$

como se requirió.

Distancia

La distancia entre dos vectores es el análogo directo de la distancia entre dos puntos en la recta numérica real o entre dos puntos en el plano cartesiano. Sobre la recta numérica (figura 1.30), la distancia entre los números a y b está dada por $|a - b|$. (Tomar el valor absoluto garantiza que no es necesario conocer cuál de a o b es mayor.) Esta distancia también es igual a $\sqrt{(a-b)^2}$, y su generalización bidimensional es la familiar fórmula para la distancia d entre los puntos (a_1, a_2) y (b_1, b_2), a saber, $d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$.

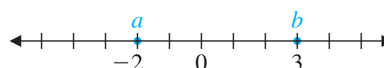


Figura 1.30
$d = |a - b| = |-2 - 3| = 5$

En términos de vectores, si $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, entonces d es justo la longitud de $\mathbf{a} - \mathbf{b}$, como se muestra en la figura 1.31. Esta es la base para la siguiente definición.

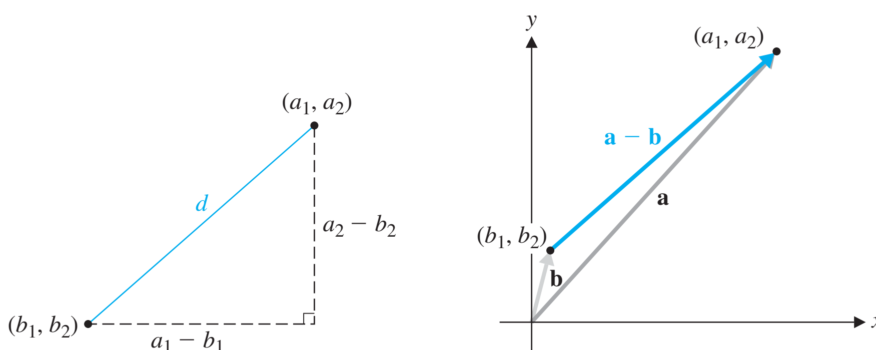


Figura 1.31
$d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2} = \|\mathbf{a} - \mathbf{b}\|$

Definición La ***distancia*** $d(\mathbf{u}, \mathbf{v})$ entre los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ se define por

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Ejemplo 1.20

Encuentre la distancia entre $\mathbf{u} = \begin{bmatrix} \sqrt{2} \\ 1 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ -2 \end{bmatrix}$.

Solución Calcule $\mathbf{u} - \mathbf{v} = \begin{bmatrix} \sqrt{2} \\ -1 \\ 1 \end{bmatrix}$, de modo que

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\sqrt{2})^2 + (-1)^2 + 1^2} = \sqrt{4} = 2$$

Ángulos

El producto punto también se puede usar para calcular el ángulo entre un par de vectores. En $\mathbb{R}^2$ o $\mathbb{R}^3$, el ángulo entre los vectores $\mathbf{u}$ y $\mathbf{v}$ distintos de cero se referirá al ángulo θ determinado por estos vectores que satisfaga $0 \le \theta \le 180°$ (vea la figura 1.32).

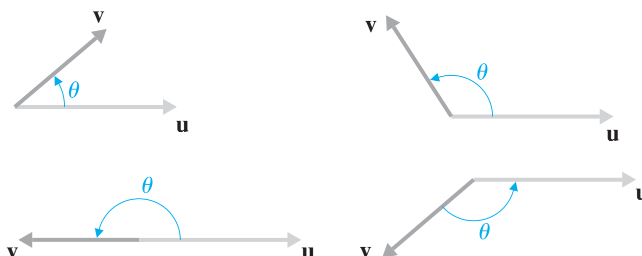


Figura 1.32
El ángulo entre **u** y **v**

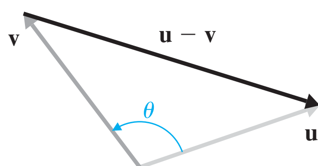


Figura 1.33

En la figura 1.33, considere el triángulo con lados **u**, **v** y **u** – **v**, donde θ es el ángulo entre **u** y **v**. Al aplicar la ley de cosenos a este triángulo se produce

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$$

Al desarrollar el lado izquierdo y usar $\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$ varias veces, se obtiene

$$\|\mathbf{u}\|^2 - 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$$

que, después de simplificar, nos deja con $\mathbf{u}\cdot\mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$. A partir de esto se obtiene la siguiente fórmula para el coseno del ángulo θ entre vectores **u** y **v** distintos de cero. Se establece como una definición.

Definición Para vectores **u** y **v** distintos de cero en $\mathbb{R}^n$,

$$\cos\theta = \frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$$

Ejemplo 1.21

Calcule el ángulo entre los vectores $\mathbf{u} = [2, 1, -2]$ y $\mathbf{v} = [1, 1, 1]$.

Solución Calcule $\mathbf{u}\cdot\mathbf{v} = 2\cdot 1 + 1\cdot 1 + (-2)\cdot 1 = 1$, $\|\mathbf{u}\| = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3$ y $\|\mathbf{v}\| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$. Por tanto, $\cos\theta = 1/3\sqrt{3}$, de modo que $\theta = \cos^{-1}(1/3\sqrt{3}) \approx 1.377$ radianes o $78.9°$.

Ejemplo 1.22

Calcule el ángulo entre las diagonales de dos caras adyacentes de un cubo.

Solución Las dimensiones del cubo no importan, así que se trabajará con un cubo con lados de longitud 1. Oriente el cubo en relación con los ejes coordenados en $\mathbb{R}^3$, como se muestra en la figura 1.34, y tome las dos diagonales laterales como los vectores [1, 0, 1] y [0, 1, 1]. Entonces el ángulo θ entre dichos vectores satisface

$$\cos\theta = \frac{1\cdot 0 + 0\cdot 1 + 1\cdot 1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}$$

de donde se concluye que el ángulo requerido es $\pi/3$ radianes o $60°$.

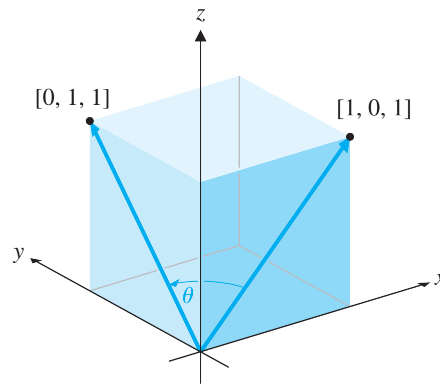


Figura 1.34

(En realidad, no es necesario hacer cálculo alguno para obtener esta respuesta. Si dibuja una tercera diagonal lateral que una los vértices en (1, 0, 1) y (0, 1, 1), obtiene un triángulo equilátero, pues todas las diagonales laterales tienen igual longitud. El ángulo que se quiere es uno de los ángulos de este triángulo y por tanto mide 60°. En ocasiones un poco de perspicacia puede ahorrar muchos cálculos; en este caso, ¡brinda una buena comprobación de su trabajo!)

Comentarios

- Como lo demuestra esta discusión, por lo general tendrá que establecer una aproximación del ángulo entre dos vectores. Sin embargo, cuando el ángulo es uno de los llamados especiales (0°, 30°, 45°, 60°, 90° o un múltiplo entero de ellos), deberá reconocer su coseno (tabla 1.1) y por ende proporcionar el ángulo correspondiente con exactitud. En todos los demás casos se usará una calculadora o computadora para aproximar el ángulo deseado mediante la función coseno inverso.

Tabla 1.1 Cosenos de ángulos especiales

θ	**0°**	**30°**	**45°**	**60°**	**90°**
$\cos\theta$	$\frac{\sqrt{4}}{2} = 1$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$	$\frac{\sqrt{1}}{2} = \frac{1}{2}$	$\frac{\sqrt{0}}{2} = 0$

- La deducción de la fórmula para el coseno del ángulo entre dos vectores es válida solamente en $\mathbb{R}^2$ y $\mathbb{R}^3$, pues depende de un hecho geométrico: la ley de los cosenos. En $\mathbb{R}^n$, para $n > 3$, la fórmula puede considerarse como una *definición*. Esto tiene sentido, pues la desigualdad de Cauchy-Schwarz implica que $\left|\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right| \leq 1$, de modo que $\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$ varía de -1 a 1, tal como hace la función coseno.

Vectores ortogonales

La palabra *ortogonal* se deriva de la palabras griegas *orthos*, que significa "derecho, recto", y *gonia*, que significa "ángulo". Por tanto, ortogonal literalmente significa "en ángulo recto". El equivalente latino es *rectangular*.

El concepto de perpendicularidad es fundamental para la geometría. Quien estudie geometría rápidamente se dará cuenta de la importancia y utilidad de los ángulos rectos. Ahora se generalizará la idea de la perpendicularidad para los vectores en $\mathbb{R}^n$, donde se le llama ***ortogonalidad***.

En $\mathbb{R}^2$ o $\mathbb{R}^3$, dos vectores $\mathbf{u}$ y $\mathbf{v}$ distintos de cero son perpendiculares si el ángulo θ entre ellos es un ángulo recto; esto es, si $\theta = \pi/2$ radianes o 90°. Por tanto, $\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|} = \cos 90° = 0$, y se concluye que $\mathbf{u}\cdot\mathbf{v} = 0$. Esto motiva la siguiente definición.

Definición Dos vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ son mutuamente ***ortogonales*** si $\mathbf{u}\cdot\mathbf{v} = 0$.

Dado que $\mathbf{0}\cdot\mathbf{v} = 0$ para todo vector $\mathbf{v}$ en $\mathbb{R}^n$, el vector cero es ortogonal a todo vector.

Ejemplo 1.23 En $\mathbb{R}^3$, $\mathbf{u} = [1, 1, -2]$ y $\mathbf{v} = [3, 1, 2]$ son ortogonales, pues $\mathbf{u}\cdot\mathbf{v} = 3 + 1 - 4 = 0$.

Al usar la noción de ortogonalidad, se obtiene una sencilla demostración del teorema de Pitágoras, válida en $\mathbb{R}^n$.

Teorema 1.6 **Teorema de Pitágoras**

Para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ si y sólo si $\mathbf{u}$ y $\mathbf{v}$ son ortogonales.

Demostración A partir del ejemplo 1.16, se tiene $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2$ para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$. Se concluye inmediatamente que $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ si y sólo si $\mathbf{u}\cdot\mathbf{v} = 0$. Vea la figura 1.35.

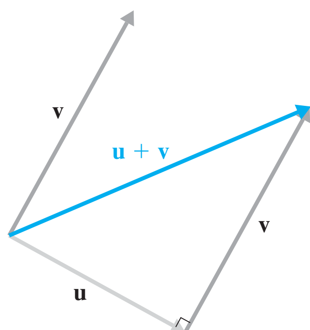


Figura 1.35

El concepto de ortogonalidad es uno de los más importantes y útiles en álgebra lineal, y con frecuencia surge en formas sorprendentes. El capítulo 5 contiene un tratamiento detallado del tema, pero se le encontrará muchas veces antes. Un problema en el que claramente tiene un papel es al encontrar la distancia desde un punto hacia una recta, donde "trazar una perpendicular" es un paso familiar.

Proyecciones

Ahora se considerará el problema de encontrar la distancia desde un punto hasta una recta en el contexto de los vectores. Como verá, esta técnica conduce a un importante concepto: la proyección de un vector sobre otro vector.

Como muestra la figura 1.36, el problema de encontrar la distancia desde un punto B hasta una recta ℓ (en $\mathbb{R}^2$ o $\mathbb{R}^3$) se reduce a encontrar la longitud del segmento de recta perpendicular $\overline{PB}$ o, de manera equivalente, la longitud del vector $\overrightarrow{PB}$. Si se elige un punto A sobre ℓ, entonces, en el triángulo recto ΔAPB, los otros dos vectores son el cateto $\overrightarrow{AP}$ y la hipotenusa $\overrightarrow{AB}$. $\overrightarrow{AP}$ se llama *proyección* de $\overrightarrow{AB}$ sobre la recta ℓ. Ahora observe esta situación en términos de vectores.

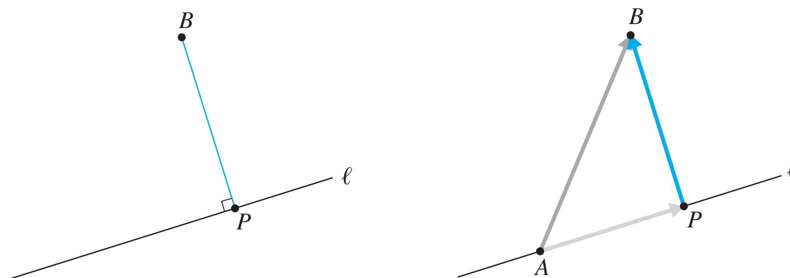


Figura 1.36
La distancia desde un punto hasta una recta

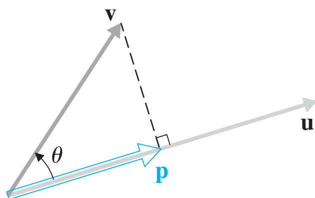


Figura 1.37
La proyección de **v** sobre **u**

Considere dos vectores **u** y **v** distintos de cero. Sea **p** el vector que se obtiene al trazar una perpendicular desde la punta de **v** sobre **u** y sea θ el ángulo entre **u** y **v**, como se muestra en la figura 1.37. Entonces, claramente $\mathbf{p} = \|\mathbf{p}\|\hat{\mathbf{u}}$, donde $\hat{\mathbf{u}} = (1/\|\mathbf{u}\|)\mathbf{u}$ es el vector unitario en la dirección de **u**. Más aún, por trigonometría elemental $\|\mathbf{p}\| = \|\mathbf{v}\| \cos\theta$, y se sabe que $\cos\theta = \dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$. Por tanto, después de sustituir, se obtiene

$$\begin{aligned}\mathbf{p} &= \|\mathbf{v}\|\left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right)\left(\frac{1}{\|\mathbf{u}\|}\right)\mathbf{u}\\ &= \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|^2}\right)\mathbf{u}\\ &= \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u}\end{aligned}$$

Ésta es la fórmula que se quería, y es la base de la siguiente definición para vectores en $\mathbb{R}^n$.

Definición Si **u** y **v** son vectores en $\mathbb{R}^n$ y $\mathbf{u} \neq \mathbf{0}$, entonces la ***proyección de* v *sobre* u** es el vector $\operatorname{proy}_{\mathbf{u}}(\mathbf{v})$ definido por

$$\operatorname{proy}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u}$$

En el ejercicio 73 se describe una forma alternativa de obtener esta fórmula.

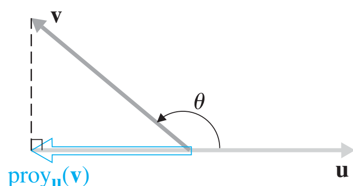


Figura 1.38

Comentarios

- El término *proyección* proviene de la idea de proyectar una imagen sobre un muro (con un proyector de diapositivas, por ejemplo). Imagine un haz de luz con rayos mutuamente paralelos y perpendiculares a **u** que brillan sobre **v**. La proyección de **v** sobre **u** es justo la sombra formada, o proyectada, por **v** sobre **u**.
- Puede ser útil considerar a $\text{proy}_{\mathbf{u}}(\mathbf{v})$ como una función con variable **v**. Entonces la variable **v** ocurre sólo una vez en el lado derecho de la definición. Además, es útil tener en mente la figura 1.38, que recuerda que $\text{proy}_{\mathbf{u}}(\mathbf{v})$ es un múltiplo escalar del vector **u** (*no* de **v**).
- Aunque en la deducción de la definición de $\text{proy}_{\mathbf{u}}(\mathbf{v})$ se requirió que tanto **v** como **u** fuesen distintos de cero (¿por qué?), es claro a partir de la geometría que la proyección del vector cero sobre **u** es **0**. La definición está en concordancia con esto, pues $\left(\dfrac{\mathbf{u}\cdot\mathbf{0}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u} = 0\mathbf{u} = \mathbf{0}$.
- Si el ángulo entre **u** y **v** es obtuso, como en la figura 1.38, entonces $\text{proy}_{\mathbf{u}}(\mathbf{v})$ estará en la dirección opuesta de **u**; esto es, $\text{proy}_{\mathbf{u}}(\mathbf{v})$ será un múltiplo escalar *negativo* de **u**.
- Si **u** es un vector unitario, entonces $\text{proy}_{\mathbf{u}}(\mathbf{v}) = (\mathbf{u}\cdot\mathbf{v})\mathbf{u}$. (¿Por qué?)

Ejemplo 1.24

Encuentre la proyección de **v** sobre **u** en cada caso.

(a) $\mathbf{v} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ (b) $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ y $\mathbf{u} = \mathbf{e}_3$

(c) $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ y $\mathbf{u} = \begin{bmatrix} 1/2 \\ 1/2 \\ 1\ \sqrt{2} \end{bmatrix}$

Solución

(a) Calcule $\mathbf{u}\cdot\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}\cdot\begin{bmatrix} -1 \\ 3 \end{bmatrix} = 1$ y $\mathbf{u}\cdot\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}\cdot\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 5$, de modo que

$$\text{proy}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u}\cdot\mathbf{v}}{\mathbf{u}\cdot\mathbf{u}}\right)\mathbf{u} = \frac{1}{5}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/5 \\ 1/5 \end{bmatrix}$$

(b) Puesto que $\mathbf{e}_3$ es un vector unitario,

$$\text{proy}_{\mathbf{e}_3}(\mathbf{v}) = (\mathbf{e}_3\cdot\mathbf{v})\mathbf{e}_3 = 3\mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

(c) Se ve que $\|\mathbf{u}\| = \sqrt{\tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{2}} = 1$. Por tanto,

$$\text{proy}_{\mathbf{u}}(\mathbf{v}) = (\mathbf{u}\cdot\mathbf{v})\mathbf{u} = \left(\frac{1}{2} + 1 + \frac{3}{\sqrt{2}}\right)\begin{bmatrix} 1/2 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix} = \frac{3(1+\sqrt{2})}{2}\begin{bmatrix} 1/2 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix}$$

$$= \frac{3(1+\sqrt{2})}{4}\begin{bmatrix} 1 \\ 1 \\ \sqrt{2} \end{bmatrix}$$

Ejercicios 1.2

En los ejercicios 1-6, encuentre $\mathbf{u} \cdot \mathbf{v}$.

1. $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ **2.** $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 9 \\ 6 \end{bmatrix}$

3. $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$ CAS **4.** $\mathbf{u} = \begin{bmatrix} 1.5 \\ 0.4 \\ -2.1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3.0 \\ 5.2 \\ -0.6 \end{bmatrix}$

5. $\mathbf{u} = [1, \sqrt{2}, \sqrt{3}, 0], \mathbf{v} = [4, -\sqrt{2}, 0, -5]$

CAS **6.** $\mathbf{u} = [1.12, -3.25, 2.07, -1.83]$,
$\mathbf{v} = [-2.29, 1.72, 4.33, -1.54]$

En los ejercicios 7-12, encuentre $\|\mathbf{u}\|$ *para el ejercicio dado y proporcione un vector unitario en la dirección de* **u**.

7. Ejercicio 1 **8.** Ejercicio 2 **9.** Ejercicio 3

CAS **10.** Ejercicio 4 **11.** Ejercicio 5 CAS **12.** Ejercicio 6

En los ejercicios 13-16, encuentre la distancia $d(\mathbf{u}, \mathbf{v})$ *entre* **u** *y* **v** *en el ejercicio dado.*

13. Ejercicio 1 **14.** Ejercicio 2

15. Ejercicio 3 CAS **16.** Ejercicio 4

17. Si **u**, **v** y **w** son vectores en $\mathbb{R}^n$, $n \geq 2$ y c es un escalar, explique por qué las siguientes expresiones no tienen sentido:

(a) $\|\mathbf{u} \cdot \mathbf{v}\|$ **(b)** $\mathbf{u} \cdot \mathbf{v} + \mathbf{w}$

(c) $\mathbf{u} \cdot (\mathbf{v} \cdot \mathbf{w})$ **(d)** $c \cdot (\mathbf{u} + \mathbf{w})$

En los ejercicios 18-23, determine si el ángulo entre **u** *y* **v** *es agudo, obtuso o recto.*

18. $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ **19.** $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$

20. $\mathbf{u} = [5, 4, -3], \mathbf{v} = [1, -2, -1]$

CAS **21.** $\mathbf{u} = [0.9, 2.1, 1.2], \mathbf{v} = [-4.5, 2.6, -0.8]$

22. $\mathbf{u} = [1, 2, 3, 4], \mathbf{v} = [-3, 1, 2, -2]$

23. $\mathbf{u} = [1, 2, 3, 4], \mathbf{v} = [5, 6, 7, 8]$

En los ejercicios 24-29, encuentre el ángulo entre **u** *y* **v** *en el ejercicio dado.*

24. Ejercico 18 **25.** Ejercicio 19 **26.** Ejercicio 20

CAS **27.** Ejercico 21 CAS **28.** Ejercicio 22 CAS **29.** Ejercicio 23

30. Sea $A = (-3, 2)$, $B = (1, 0)$ y $C = (4, 6)$. Pruebe que ΔABC es un triángulo recto.

31. Sea $A = (1, 1, -1)$, $B = (-3, 2, -2)$, y $C = (2, 2, -4)$. Pruebe que ΔABC es un triángulo rectángulo.

CAS **32.** Encuentre el ángulo entre una diagonal de un cubo y una arista adyacente.

33. Un cubo tiene cuatro diagonales. Demuestre que ningún par de ellas son perpendiculares.

En los ejercicios 34-39, encuentre la proyección de **v** *sobre* **u**. *Dibuje un bosquejo para los ejercicios 34 y 35.*

34. Un paralelogramo tiene diagonales determinadas por los vectores

$$\mathbf{d}_1 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{d}_2 = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}$$

Demuestre que el paralelogramo es un rombo (todos los lados de igual longitud) y determine la longitud del lado.

35. El rectángulo $ABCD$ tiene vértices en $A = (1, 2, 3)$, $B = (3, 6, -2)$ y $C = (0, 5, -4)$. Determine las coordenadas del vértice D.

36. Un avión se dirige hacia el este con una velocidad de 200 millas por hora. Un viento sopla del norte a 40 millas por hora. ¿Cuál es la velocidad resultante del avión?

37. Un bote se dirige al norte a través de un río con una velocidad de 4 millas por hora. Si la corriente fluye al este con una velocidad de 3 millas por hora, encuentre la velocidad resultante del bote.

38. Ann conduce un bote de motor a través de un río que tiene 2 km de ancho. El bote tiene una rapidez de 20 km/h en aguas quietas, y la corriente en el río fluye a 5 km/h. Ann parte desde una orilla en el río hacia un muelle directamente frente a ella en la ribera opuesta. Ella conduce el bote en una dirección perpendicular a la corriente.

(a) ¿Cuán lejos río abajo del muelle llegará Ann?
(b) ¿Cuánto tardará Ann en cruzar el río?

39. Bert puede nadar a un ritmo de 2 millas por hora en aguas quietas. La corriente en un río fluye con una velocidad de 1 milla por hora. Si Bert quiere nadar a través del río hasta un punto directamente opuesto, ¿a qué ángulo con la orilla del río debe nadar?

En los ejercicios 40-45, encuentre la proyección de **v** *sobre* **u**. *Dibuje un bosquejo para los ejercicios 40 y 41.*

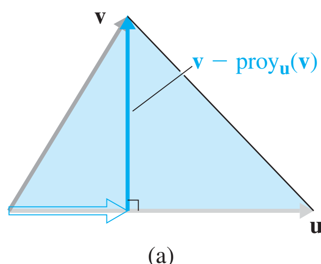
40. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ -1 \end{bmatrix}$ **41.** $\mathbf{u} = \begin{bmatrix} 3/5 \\ -4/5 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

42. $\mathbf{u} = \begin{bmatrix} 2/3 \\ -2/3 \\ -1/3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$ **43.** $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -3 \\ -1 \\ -2 \end{bmatrix}$

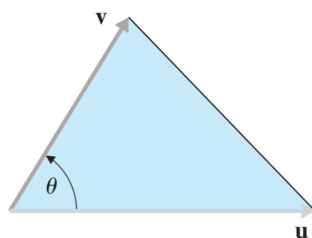
CAS **44.** $\mathbf{u} = \begin{bmatrix} 0.5 \\ 1.5 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2.1 \\ 1.2 \end{bmatrix}$

CAS **45.** $\mathbf{u} = \begin{bmatrix} 3.01 \\ -0.33 \\ 2.52 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1.34 \\ 4.25 \\ -1.66 \end{bmatrix}$

La figura 1.39 sugiere dos formas en las que pueden usarse los vectores para calcular el área de un triángulo. El área $\mathcal{A}$ del triángulo en el inciso (a) está dada por $\frac{1}{2}\|\mathbf{u}\|\|\mathbf{v} - \text{proy}_{\mathbf{u}}(\mathbf{v})\|$, y el inciso (b) sugiere la forma trigonométrica del área de un triángulo:



(a)



(b)

Figura 1.39

$\mathcal{A} = \frac{1}{2}\|\mathbf{u}\|\|\mathbf{v}\|\operatorname{sen}\theta$. (Puede usar la identidad $\operatorname{sen}\theta = \sqrt{1 - \cos^2\theta}$ para encontrar sen θ.)

En los ejercicios 46 y 47, calcule el área del triángulo con los vértices dados usando ambos métodos.

46. $A = (1, -1), B = (2, 2), C = (4, 0)$

47. $A = (3, -1, 4), B = (4, -2, 6), C = (5, 0, 2)$

En los ejercicios 48 y 49, encuentre todos los valores del escalar k para el cual los dos vectores son ortogonales.

48. $\mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} k+1 \\ k-1 \end{bmatrix}$ **49.** $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} k^2 \\ k \\ -3 \end{bmatrix}$

50. Describa todos los vectores $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ que son ortogonales a $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$.

51. Describa todos los vectores $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ que son ortogonales a $\mathbf{u} = \begin{bmatrix} a \\ b \end{bmatrix}$.

52. En qué condiciones los siguientes son verdaderos para vectores **u** y **v** en $\mathbb{R}^2$ o $\mathbb{R}^3$?

(a) $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$ **(b)** $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| - \|\mathbf{v}\|$

53. Demuestre el Teorema 1.2(b).

54. Demuestre el Teorema 1.2(d).

En los ejercicios 55-57, demuestre la propiedad enunciada de la distancia entre vectores.

55. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$ para todos los vectores **u** y **v**.

56. $d(\mathbf{u}, \mathbf{w}) \le d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w})$ para todos los vectores **u**, **v** y **w**

57. $d(\mathbf{u}, \mathbf{v}) = 0$ si y sólo si $\mathbf{u} = \mathbf{v}$

58. Demuestre que $\mathbf{u} \cdot c\mathbf{v} = c(\mathbf{u} \cdot \mathbf{v})$ para todos los vectores **u** y **v** en $\mathbb{R}^n$ y todo escalar c.

59. Demuestre que $\|\mathbf{u} - \mathbf{v}\| \ge \|\mathbf{u}\| - \|\mathbf{v}\|$ para todos los vectores **u** y **v** en $\mathbb{R}^n$. [*Sugerencia:* sustituya **u** por $\mathbf{u} - \mathbf{v}$ en la desigualdad del triángulo.]

60. Suponga que usted sabe que $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$. ¿Se deduce que $\mathbf{v} = \mathbf{w}$? Si es así, proporcione una demostración que sea válida en $\mathbb{R}^n$; de otro modo, proporcione un *contraejemplo* (esto es, un conjunto *específico* de vectores **u**, **v** y **w** para los cuales $\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot \mathbf{w}$ pero $\mathbf{v} \ne \mathbf{w}$).

61. Demuestre que $(\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v}) = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2$ para todos los vectores **u** y **v** en $\mathbb{R}^n$.

62. **(a)** Demuestre que $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$ para todos los vectores **u** y **v** en $\mathbb{R}^n$.

(b) Dibuje un diagrama que muestre **u**, **v**, $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ en $\mathbb{R}^2$ y use (a) para deducir un resultado acerca de los paralelogramos.

63. Demuestre que $\mathbf{u} \cdot \mathbf{v} = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$ para todos los vectores **u** y **v** en $\mathbb{R}^n$.

64. **(a)** Demuestre que $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\|$ si y sólo si **u** y **v** son ortogonales.
(b) Dibuje un diagrama que muestre **u**, **v**, **u** + **v** y **u** – **v** en $\mathbb{R}^2$ y use (a) para deducir un resultado acerca de los paralelogramos.

65. **(a)** Demuestre que **u** + **v** y **u** – **v** son ortogonales en $\mathbb{R}^n$ si y sólo si $\|\mathbf{u}\| = \|\mathbf{v}\|$.
(b) Dibuje un diagrama que muestre **u**, **v**, **u** + **v** y **u** – **v** en $\mathbb{R}^2$ y use (a) para deducir un resultado acerca de los paralelogramos.

66. Si $\|\mathbf{u}\| = 2$, $\|\mathbf{v}\| = \sqrt{3}$, y $\mathbf{u} \cdot \mathbf{v} = 1$, calcule $\|\mathbf{u} + \mathbf{v}\|$.

67. Demuestre que no hay vectores **u** y **v** tales que $\|\mathbf{u}\| = 1$, $\|\mathbf{v}\| = 2$ y $\mathbf{u} \cdot \mathbf{v} = 3$.

68. **(a)** Demuestre que si **u** es ortogonal tanto a **v** como a **w**, entonces **u** es ortogonal a **v** + **w**.
(b) Demuestre que si **u** es ortogonal tanto a **v** como a **w**, entonces **u** es ortogonal a $s\mathbf{v} + t\mathbf{w}$ para todos los escalares s y t.

69. Demuestre que **u** es ortogonal a $\mathbf{v} - \text{proy}_{\mathbf{u}}(\mathbf{v})$ para todos los vectores **u** y **v** en $\mathbb{R}^n$, donde $\mathbf{u} \neq \mathbf{0}$.

70. **(a)** Demuestre que $\text{proy}_{\mathbf{u}}(\text{proy}_{\mathbf{u}}(\mathbf{v})) = \text{proy}_{\mathbf{u}}(\mathbf{v})$.
(b) Demuestre que $\text{proy}_{\mathbf{u}}(\mathbf{v} - \text{proy}_{\mathbf{u}}(\mathbf{v})) = \mathbf{0}$.
(c) Explique (a) y (b) geométricamente.

71. La desigualdad de Cauchy-Schwarz $|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \|\mathbf{v}\|$ es equivalente a la desigualdad que se obtiene al elevar al cuadrado ambos lados: $(\mathbf{u} \cdot \mathbf{v})^2 \leq \|\mathbf{u}\|^2 \|\mathbf{v}\|^2$.
(a) En $\mathbb{R}^2$ con $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, esto se convierte en

$$(u_1v_1 + u_2v_2)^2 \leq (u_1^2 + u_2^2)(v_1^2 + v_2^2)$$

Demuestre esto algebraicamente. [*Sugerencia:* reste el lado izquierdo del lado derecho y demuestre que la diferencia necesariamente debe ser no negativa.]
(b) Demuestre el análogo de (a) en $\mathbb{R}^3$.

72. En la figura 1.40 se sugiere otro método para demostrar la desigualdad de Cauchy-Schwarz, donde se muestra que, en $\mathbb{R}^2$ o $\mathbb{R}^3$, $\|\text{proy}_{\mathbf{u}}(\mathbf{v})\| \leq \|\mathbf{v}\|$. Demuestre que esto es equivalente a la desigualdad de Cauchy-Schwarz.

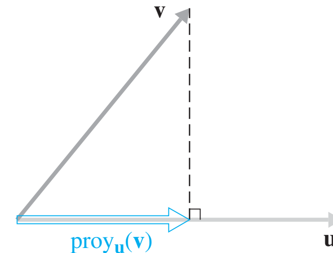


Figura 1.40

73. Use el hecho de que $\text{proy}_{\mathbf{u}}(\mathbf{v}) = c\mathbf{u}$ para algún escalar c, junto con la figura 1.41, para encontrar c y con ello deducir la fórmula para $\text{proy}_{\mathbf{u}}(\mathbf{v})$.

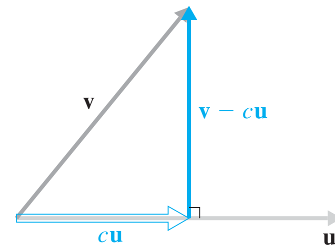


Figura 1.41

74. Con inducción matemática, demuestre la siguiente generalización de la desigualdad del triángulo:

$$\|\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n\| \leq \|\mathbf{v}_1\| + \|\mathbf{v}_2\| + \cdots + \|\mathbf{v}_n\|$$

para todo $n \geq 1$.

Exploración

Vectores y geometría

Muchos resultados en geometría euclideana plana pueden demostrarse con el uso de técnicas vectoriales. Por ejemplo, en el ejemplo 1.24 se usaron vectores para demostrar el teorema de Pitágoras. En esta exploración se usarán vectores para desarrollar demostraciones para algunos otros teoremas de la geometría euclideana.

Como introducción a la notación y el procedimiento básico, considere el siguiente ejemplo sencillo.

Ejemplo 1.25

Proporcione una descripción vectorial del punto medio M de un segmento de recta $\overline{AB}$.

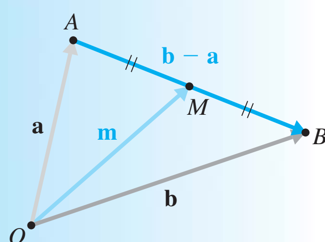


Figura 1.42
El punto medio de $\overline{AB}$

Solución Primero convierta todo a notación vectorial. Si O denota el origen y P es un punto, sea $\mathbf{p}$ el vector $\overrightarrow{OP}$. En esta situación, $\mathbf{a} = \overrightarrow{OA}$, $\mathbf{b} = \overrightarrow{OB}$, $\mathbf{m} = \overrightarrow{OM}$ y $\overrightarrow{AB} = \overrightarrow{OB} - \overrightarrow{OA} = \mathbf{b} - \mathbf{a}$ (figura 1.42).

Ahora, dado que M es el punto medio de $\overline{AB}$, se tiene

$$\mathbf{m} - \mathbf{a} = \overrightarrow{AM} = \tfrac{1}{2}\overrightarrow{AB} = \tfrac{1}{2}(\mathbf{b} - \mathbf{a})$$

de modo que

$$\mathbf{m} = \mathbf{a} + \tfrac{1}{2}(\mathbf{b} - \mathbf{a}) = \tfrac{1}{2}(\mathbf{a} + \mathbf{b})$$

1. Proporcione una descripción vectorial del punto P que está a un tercio del camino entre A y B sobre el segmento de recta $\overline{AB}$. Generalice.

2. Pruebe que el segmento de recta que une los puntos medios de dos lados de un triángulo es paralelo al tercer lado y tiene la mitad de la longitud. (En notación vectorial, pruebe que $\overrightarrow{PQ} = \frac{1}{2}\overrightarrow{AB}$ en la figura 1.43.)

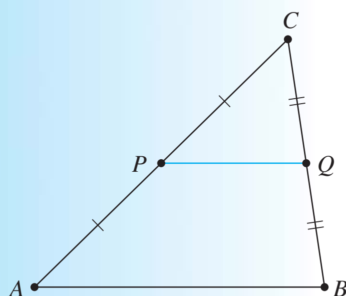


Figura 1.43

3. Pruebe que el cuadrilátero $PQRS$ (figura 1.44), cuyos vértices están en los puntos medios de los lados de un cuadrilátero arbitrario $ABCD$, es un paralelogramo.

4. Una ***mediana*** de un triángulo es un segmento de recta desde un vértice hasta el punto medio del lado opuesto (figura 1.45). Pruebe que las tres medianas de cualquier triángulo son *concurrentes* (es decir: tienen un punto de intersección común) en un punto G que está a dos tercios de la distancia desde cada vértice hasta el punto medio del lado opuesto. [*Sugerencia:* en la figura 1.46 se muestra que el punto que está a dos tercios de la distancia de A a P está dado por $\frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c})$. Entonces demuestre que $\frac{1}{3}(\mathbf{a} + \mathbf{b} + \mathbf{c})$ es dos tercios la distancia de B a Q y dos tercios la distancia de C a R.] El punto G en la figura 1.46 se llama ***centroide*** del triángulo.

5. La ***altura*** de un triángulo es un segmento de recta desde un vértice y que es perpendicular al lado opuesto (figura 1.47). Demuestre que las tres alturas de un triángulo son concurrentes. [*Sugerencia:* sea H el punto de intersección de las alturas desde A y B

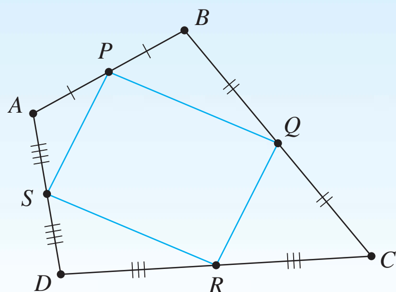


Figura 1.44

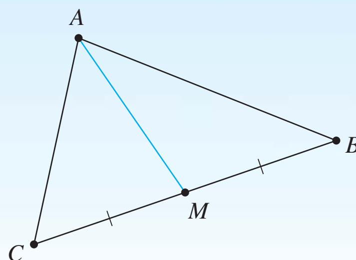


Figura 1.45
Una mediana

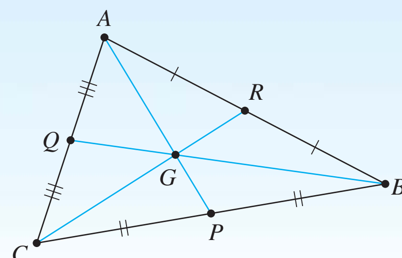
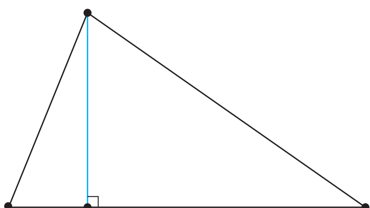


Figura 1.46
El centroide

en la figura 1.48. Pruebe que $\overrightarrow{CH}$ es ortogonal a $\overrightarrow{AB}$.] El punto H en la figura 1.48 se llama ***ortocentro*** del triángulo.

6. Un ***bisector perpendicular*** de un segmento de recta es una recta que pasa por el punto medio del segmento, perpendicular a él (figura 1.49). Demuestre que los bisectores perpendiculares de los tres lados de un triángulo son concurrentes. [*Sugerencia:* sea K el punto de intersección de los bisectores perpendiculares de $\overline{AC}$ y $\overline{BC}$ en la figura 1.50. Demuestre que $\overrightarrow{RK}$ es ortogonal a $\overrightarrow{AB}$.] El punto K en la figura 1.50 se conoce como ***circuncentro*** del triángulo.


Figura 1.47
Una altura

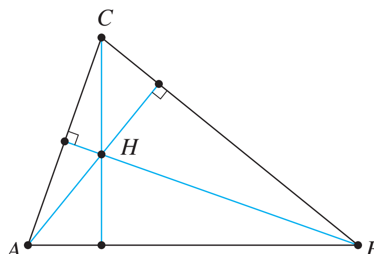


Figura 1.48
El ortocentro

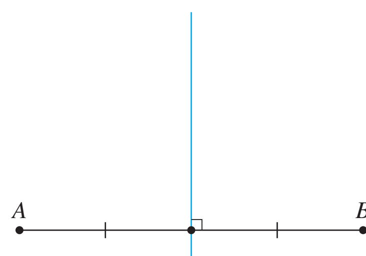


Figura 1.49
Un bisector perpendicular

7. Sean A y B los puntos extremos de un diámetro de un círculo. Si C es cualquier punto sobre la circunferencia, demuestre que $\angle ACB$ es un ángulo recto. [*Sugerencia:* en la figura 1.51, sea O el centro del círculo. Exprese todo en términos de **a** y **c** y demuestre que $\overrightarrow{AC}$ es ortogonal a $\overrightarrow{BC}$.]

8. Demuestre que los segmentos de recta que unen los puntos medios de lados opuestos de un cuadrilátero se bisecan mutuamente (figura 1.52).

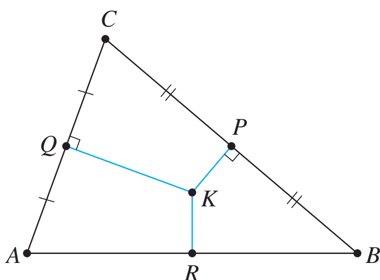


Figura 1.50
El circuncentro

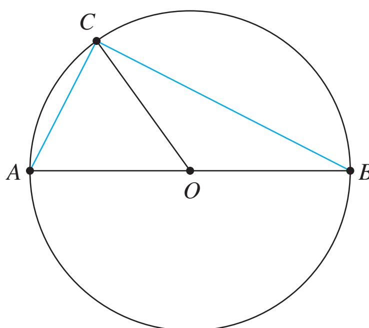


Figura 1.51

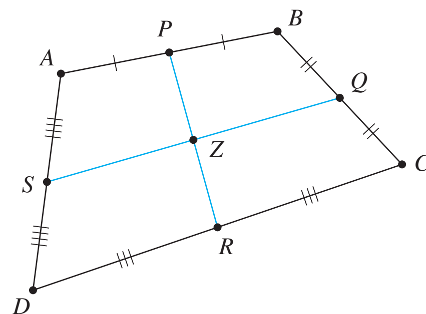


Figura 1.52

1.3 Rectas y planos

Todos están familiarizados con la ecuación de una recta en el plano cartesiano. Ahora se considerarán rectas en $\mathbb{R}^2$ desde un punto de vista vectorial. La comprensión que se obtenga a partir de este planteamiento permitirá generalizar a rectas en $\mathbb{R}^3$ y luego a planos en $\mathbb{R}^3$. Mucha del álgebra lineal que se considerará en capítulos posteriores tiene sus orígenes en la geometría simple de rectas y planos; la habilidad para visualizarlos y pensar geométricamente en torno a un problema le servirá bastante.

Rectas en $\mathbb{R}^2$ y $\mathbb{R}^3$

En el plano xy, la forma general de la ecuación de una recta es $ax + by = c$. Si $b \neq 0$, entonces la ecuación puede reescribirse como $y = -(a/b)x + c/b$, que tiene la forma $y = mx + k$. [Ésta es la forma pendiente ordenada al origen; m es la pendiente de la recta y el punto con coordenadas $(0, k)$ es su ordenada al origen.] Para incluir los vectores en este estudio, considere un ejemplo.

Ejemplo 1.26

En la figura 1.53 se muestra la recta ℓ, con ecuación $2x + y = 0$. Es una recta con pendiente -2 que pasa por el origen. El lado izquierdo de la ecuación está en la forma de producto punto; de hecho, si $\mathbf{n} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, entonces la ecuación se convierte en $\mathbf{n} \cdot \mathbf{x} = 0$. El vector $\mathbf{n}$ es perpendicular a la recta; esto es, es *ortogonal* a cualquier vector $\mathbf{x}$ que sea paralelo a la recta (figura 1.54) y se le conoce como ***vector normal*** a la recta. La ecuación $\mathbf{n} \cdot \mathbf{x} = 0$ es la *forma normal* de la ecuación de ℓ.

Otra forma de pensar esta recta es imaginar una partícula que se mueve a lo largo de la recta. Suponga que la partícula inicialmente está en el origen en el tiempo $t = 0$ y se mueve a lo largo de la recta en tal forma que su coordenada x cambia 1 unidad por segundo. Entonces, en $t = 1$ la partícula está en $(1, -2)$, en $t = 1.5$ está en $(1.5, -3)$ y si se permiten valores negativos de t (esto es, considera dónde estuvo la partícula en el pasado), en $t = -2$ está (o estuvo) en $(-2, 4)$. Este movimiento se ilustra en la figura 1.55.

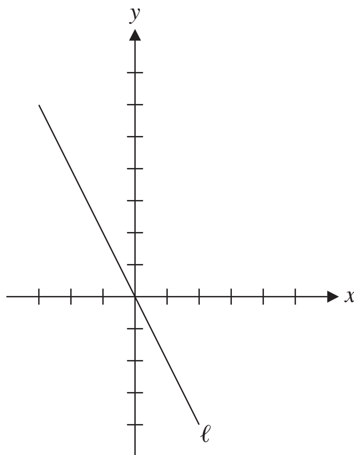


Figura 1.53
La recta $2x + y = 0$

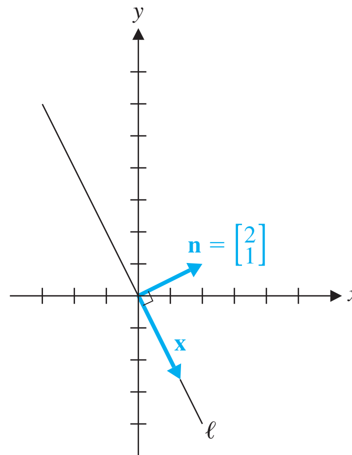
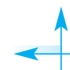


Figura 1.54
Un vector normal $\mathbf{n}$

La palabra latina *norma* se refiere a una escuadra de carpintero que se usaba para dibujar ángulos rectos. Por ende, un vector **normal** es aquel que es perpendicular a algo más, por lo general un plano.



En general, si $x = t$, entonces $y = -2t$, y esta relación se puede escribir en forma vectorial como

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} t \\ -2t \end{bmatrix} = t\begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

¿Cuál es el significado del vector $\mathbf{d} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$? Es un vector particular paralelo a ℓ, llamado ***vector director*** para la recta. Como se muestra en la figura 1.56, puede escribir la ecuación de ℓ como $\mathbf{x} = t\mathbf{d}$. Esta es la *forma vectorial* de la ecuación de la recta.

Si la recta no pasa por el origen, entonces deben modificarse ligeramente las cosas.

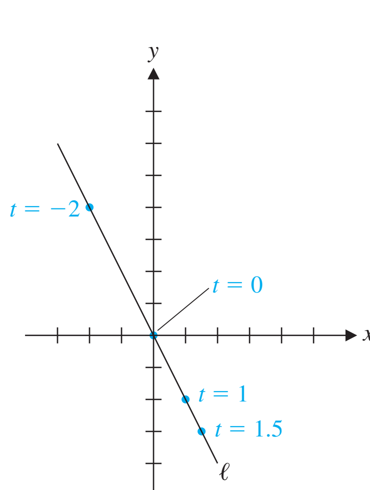


Figura 1.55

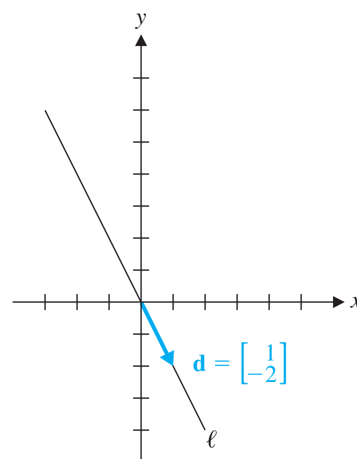


Figura 1.56
Un vector de dirección **d**

Ejemplo 1.27

Considere la recta ℓ con ecuación $2x + y = 5$ (figura 1.57). Esta es justo la recta del ejemplo 1.26 desplazada 5 unidades arriba. También tiene pendiente -2, pero su ordenada al origen es el punto $(0, 5)$. Es claro que los vectores **d** y **n** del ejemplo 1.26 son, respectivamente, un vector director y un vector normal también a esta recta.

Por ende, **n** es ortogonal a cada vector que sea paralelo a ℓ. El punto $P = (1, 3)$ está sobre ℓ. Si $X = (x, y)$ representa un punto general sobre ℓ, entonces el vector $\overrightarrow{PX} = \mathbf{x} - \mathbf{p}$ es paralelo a ℓ y $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$ (vea la figura 1.58). Al simplificar, se tiene $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$. Como comprobación, calcule

$$\mathbf{n} \cdot \mathbf{x} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 2x + y \quad \text{y} \quad \mathbf{n} \cdot \mathbf{p} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 5$$

Por tanto, la forma normal $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$ sólo es una representación diferente de la forma general de la ecuación de la recta. (Note que en el ejemplo 1.26, **p** fue el vector cero, de modo que $\mathbf{n} \cdot \mathbf{p} = 0$ produjo el lado derecho de la ecuación.)

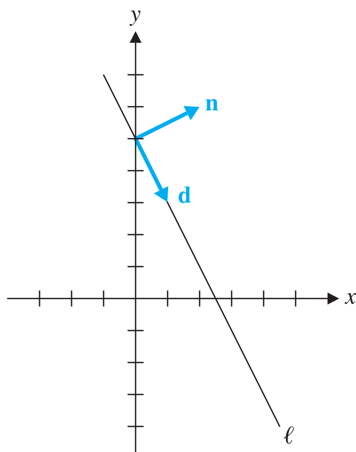


Figura 1.57
La recta $2x + y = 5$

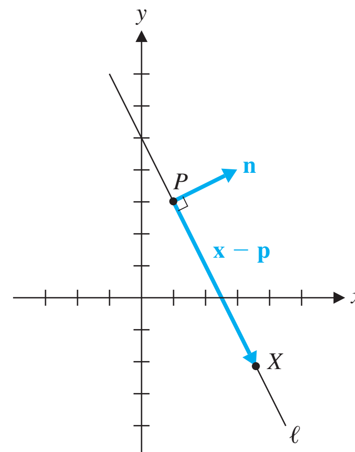


Figura 1.58
$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$

Estos resultados conducen a la siguiente definición.

Definición La ***forma normal de la ecuación de una recta*** ℓ en $\mathbb{R}^2$ es

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0 \quad \text{o} \quad \mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$$

donde $\mathbf{p}$ es un punto específico sobre ℓ y $\mathbf{n} \neq \mathbf{0}$ es un vector normal a ℓ.

La ***forma general de la ecuación de*** ℓ es $ax + by = c$, donde $\mathbf{n} = \begin{bmatrix} a \\ b \end{bmatrix}$ es un vector normal para ℓ.

Continuando con el ejemplo 1.27, ahora se encontrará la forma vectorial de la ecuación de ℓ. Note que, para cada elección de $\mathbf{x}$, $\mathbf{x} - \mathbf{p}$ debe ser paralelo a, y por tanto múltiplo de, el vector director $\mathbf{d}$. Esto es $\mathbf{x} - \mathbf{p} = t\mathbf{d}$ o $\mathbf{x} = \mathbf{p} + t\mathbf{d}$ para algún escalar t. En términos de componentes, se tiene

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \end{bmatrix} \tag{1}$$

o

$$\begin{aligned} x &= 1 + t \\ y &= 3 - 2t \end{aligned} \tag{2}$$

La ecuación (1) es la forma vectorial de la ecuación de ℓ, y las ecuaciones en componentes (2) se llaman *ecuaciones paramétricas* de la recta. La variable t se conoce como *parámetro*.

La palabra *parámetro*, y el correspondiente adjetivo *paramétrica*, provienen de las palabras griegas *para*, que significa "a lo largo de, junto a", y *metron*, que significa "medir". En lenguaje matemático, un parámetro es una variable en términos de la cual se expresan otras variables: una nueva "medida" se coloca junto a las anteriores.

¿Cómo se generaliza todo esto a $\mathbb{R}^3$? Observe que las formas vectorial y paramétrica de las ecuaciones de una recta se trasladan perfectamente. La noción de la pendiente de una recta en $\mathbb{R}^2$, que es difícil de generalizar a tres dimensiones, se sustituye por la noción más conveniente de vector director, lo que conduce a la siguiente definición.

Definición La ***forma vectorial de la ecuación de una recta*** ℓ en $\mathbb{R}^2$ o $\mathbb{R}^3$ es

$$\mathbf{x} = \mathbf{p} + t\mathbf{d}$$

donde $\mathbf{p}$ es un punto específico sobre ℓ y $\mathbf{d} \neq \mathbf{0}$ es un vector director para ℓ.

Las ecuaciones que corresponden a los componentes de la forma vectorial de la ecuación se llaman ***ecuaciones paramétricas*** de ℓ.

Con frecuencia, esta terminología se abreviará ligeramente, y se hará referencia simplemente a las ecuaciones general, normal, vectorial y paramétrica de una recta o plano.

Ejemplo 1.28 Encuentre las ecuaciones vectorial y paramétrica de la recta en $\mathbb{R}^3$ que pasa por el punto $P = (1, 2, -1)$, paralela al vector $\mathbf{d} = \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$.

Solución La ecuación vectorial $\mathbf{x} = \mathbf{p} + t\mathbf{d}$ es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + t\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$$

La forma paramétrica es

$$\begin{aligned} x &= 1 + 5t \\ y &= 2 - t \\ z &= -1 + 3t \end{aligned}$$

Comentarios

- Las formas vectorial y paramétrica de la ecuación de una recta dada ℓ no son únicas; de hecho, existen infinitas formas, pues es posible usar cualquier punto sobre ℓ para determinar $\mathbf{p}$ y cualquier vector director para ℓ. Sin embargo, claramente todos los vectores directores son múltiplos mutuos.

En el ejemplo 1.28, $(6, 1, 2)$ es otro punto sobre la recta (considere $t = 1$) y $\begin{bmatrix} 10 \\ -2 \\ 6 \end{bmatrix}$ es otro vector director. Por tanto,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ 2 \end{bmatrix} + s\begin{bmatrix} 10 \\ -2 \\ 6 \end{bmatrix}$$

proporciona una ecuación vectorial diferente (pero equivalente) para la recta. La relación entre los dos parámetros s y t puede encontrarse al comparar las ecuaciones paramétricas: para un punto dado (x, y, z) sobre ℓ, se tiene

$$\begin{aligned} x &= 1 + 5t = 6 + 10s \\ y &= 2 - t = 1 - 2s \\ z &= -1 + 3t = 2 + 6s \end{aligned}$$

lo que implica que

$$\begin{aligned} -10s + 5t &= 5 \\ 2s - t &= -1 \\ -6s + 3t &= 3 \end{aligned}$$

Cada una de dichas ecuaciones se reduce a $t = 1 + 2s$.

- Intuitivamente se sabe que una recta es un objeto *unidimensional*. La idea de "dimensión" se aclarará en los capítulos 3 y 6, pero por el momento observe que esta idea parece concordar con el hecho de que la forma vectorial de la ecuación de una recta requiere *un* parámetro.

Ejemplo 1.29

Con frecuencia, uno escucha la expresión "dos puntos determinan una recta". Encuentre una ecuación vectorial de la recta ℓ en $\mathbb{R}^3$ determinada por los puntos $P = (-1, 5, 0)$ y $Q = (2, 1, 1)$.

Solución Es posible elegir cualquier punto sobre ℓ para $\mathbf{p}$, de modo que se usará P (Q también estaría bien).

Un vector director conveniente es $\mathbf{d} = \overrightarrow{PQ} = \begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix}$ (o cualquier múltiplo escalar de éste). Por tanto, se obtiene

$$\begin{aligned} \mathbf{x} &= \mathbf{p} + t\mathbf{d} \\ &= \begin{bmatrix} -1 \\ 5 \\ 0 \end{bmatrix} + t\begin{bmatrix} 3 \\ -4 \\ 1 \end{bmatrix} \end{aligned}$$

Planos en $\mathbb{R}^3$

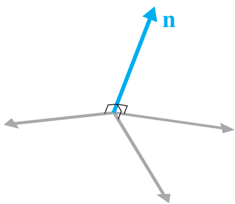


Figura 1.59
$\mathbf{n}$ es ortogonal a un número infinito de vectores

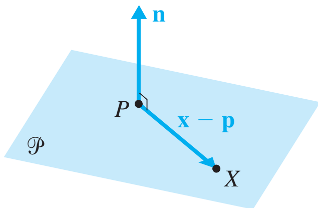


Figura 1.60
$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$

La siguiente pregunta que debe plantearse es: ¿cómo la forma general de la ecuación de una recta se generaliza a $\mathbb{R}^3$? Razonablemente puede suponer que si $ax + by = c$ es la forma general de la ecuación de una recta en $\mathbb{R}^2$, entonces $ax + by + cz = d$ puede representar una recta en $\mathbb{R}^3$. En forma normal, esta ecuación sería $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$, donde $\mathbf{n}$ es un vector normal a la recta y $\mathbf{p}$ corresponde a un punto sobre la recta.

Para ver si esta es una hipótesis razonable, considere el caso especial de la ecuación $ax + by + cz = 0$. En forma normal, se convierte en $\mathbf{n} \cdot \mathbf{x} = 0$, donde $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$. Sin embargo, el conjunto de todos los vectores $\mathbf{x}$ que satisfacen esta ecuación es el conjunto de todos los vectores ortogonales a $\mathbf{n}$. Como se muestra en la figura 1.59, vectores en infinitas direcciones tienen esta propiedad, lo que determina una familia de *planos* paralelos. De modo que la suposición fue incorrecta: parece que $ax + by + cz = d$ es la ecuación de un plano, no de una recta, en $\mathbb{R}^3$.

Precise más este hallazgo. Todo plano $\mathcal{P}$ en $\mathbb{R}^3$ puede determinarse al especificar un punto $\mathbf{p}$ sobre $\mathcal{P}$ y un vector distinto de cero $\mathbf{n}$ normal a $\mathcal{P}$ (figura 1.60). Por ende, si $\mathbf{x}$ representa un punto arbitrario sobre $\mathcal{P}$, se tiene que $\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0$ o $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$. si $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$, entonces, en términos de componentes, la ecuación se convierte en $ax + by + cz = d$ (donde $d = \mathbf{n} \cdot \mathbf{p}$).

Definición La ***forma normal de la ecuación de un plano*** $\mathcal{P}$ en $\mathbb{R}^3$ es

$$\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}) = 0 \quad \text{o} \quad \mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$$

donde $\mathbf{p}$ es un punto específico sobre $\mathcal{P}$ y $\mathbf{n} \neq \mathbf{0}$ es un vector normal para $\mathcal{P}$.

La ***forma general de la ecuación de*** $\mathcal{P}$ es $ax + by + cz = d$, donde $\mathbf{n} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ es un vector normal para $\mathcal{P}$.

Note que cualquier múltiplo escalar de un vector normal para un plano es otro vector normal.

Ejemplo 1.30

Encuentre las formas normal y general de la ecuación del plano que contiene el punto $P = (6, 0, 1)$ y tiene un vector normal $\mathbf{n} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Solución Con $\mathbf{p} = \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ se tiene $\mathbf{n} \cdot \mathbf{p} = 1 \cdot 6 + 2 \cdot 0 + 3 \cdot 1 = 9$, de modo que la ecuación normal $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$ se convierte en la ecuación general $x + 2y + 3z = 9$.

Geométricamente, es claro que planos paralelos tienen los mismos vectores normales. Por tanto, sus ecuaciones generales tienen lados izquierdos que son múltiplos mutuos. De este modo, por ejemplo, $2x + 4y + 6z = 10$ es la ecuación general de un plano que es paralelo al plano del ejemplo 1.30, ya que la ecuación puede reescribirse como $x + 2y + 3z = 5$, de donde se ve que los dos planos tienen el mismo vector normal $\mathbf{n}$. (Note que los planos no coinciden, pues los lados derechos de sus ecuaciones son distintos.)

También puede expresar la ecuación de un plano en forma vectorial o paramétrica. Para hacerlo, observe que un plano también puede determinarse al especificar uno de sus puntos P (por el vector $\mathbf{p}$) y *dos* vectores directores $\mathbf{u}$ y $\mathbf{v}$ paralelos al plano (pero no mutuamente paralelos). Como muestra la figura 1.61, dado cualquier punto X en el plano (ubicado por $\mathbf{x}$), siempre se pueden encontrar múltiplos adecuados $s\mathbf{u}$ y $t\mathbf{v}$ de los vectores directores tales que $\mathbf{x} - \mathbf{p} = s\mathbf{u} + t\mathbf{v}$ or $\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$. Si esta ecuación se escribe en forma de componentes, se obtienen ecuaciones paramétricas para el plano.

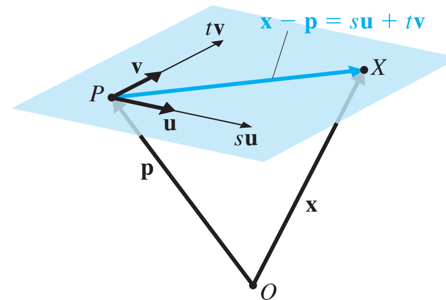


Figura 1.61
$\mathbf{x} - \mathbf{p} = s\mathbf{u} + t\mathbf{v}$

Definición La ***forma vectorial de la ecuación de un plano*** $\mathcal{P}$ en $\mathbb{R}^3$ es

$$\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$$

donde $\mathbf{p}$ es un punto sobre $\mathcal{P}$ y $\mathbf{u}$, y $\mathbf{v}$ son vectores directores para $\mathcal{P}$ ($\mathbf{u}$ y $\mathbf{v}$ son distintos de cero y paralelos a $\mathcal{P}$, mas no mutuamente paralelos).

Las ecuaciones correspondientes a los componentes de la forma vectorial de la ecuación se llaman ***ecuaciones paramétricas*** de $\mathcal{P}$.

Ejemplo 1.31

Encuentre ecuaciones vectoriales y paramétricas para el plano del ejemplo 1.30.

Solución Es necesario encontrar dos vectores directores. Se tiene un punto, $P = (6, 0, 1)$, en el plano; si puede encontrar otros dos puntos Q y R en $\mathcal{P}$, entonces los vectores $\overrightarrow{PQ}$ y $\overrightarrow{PR}$ pueden servir como vectores directores (a menos que, por mala suerte, ¡resulten ser paralelos!). Mediante ensayo y error se observa que $Q = (9, 0, 0)$ y $R = (3, 3, 0)$ satisfacen la ecuación general $x + 2y + 3z = 9$ y por tanto yacen en el plano. Entonces se calcula

$$\mathbf{u} = \overrightarrow{PQ} = \mathbf{q} - \mathbf{p} = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \overrightarrow{PR} = \mathbf{r} - \mathbf{p} = \begin{bmatrix} -3 \\ 3 \\ -1 \end{bmatrix}$$

que, dado que no son múltiplos escalares mutuos, servirán como vectores directores. Por tanto, se tiene la ecuación vectorial de $\mathcal{P}$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix} + s\begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} + t\begin{bmatrix} -3 \\ 3 \\ -1 \end{bmatrix}$$

y las correspondientes ecuaciones paramétricas,

$$\begin{aligned} x &= 6 + 3s - 3t \\ y &= 3t \\ z &= 1 - s - t \end{aligned}$$

[¿Qué habría ocurrido si hubiera escogido $R = (0, 0, 3)$?]

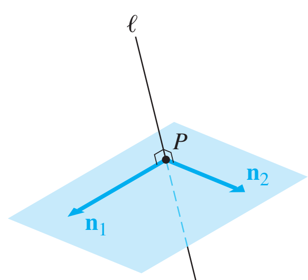


Figura 1.62
Dos normales determinan una recta

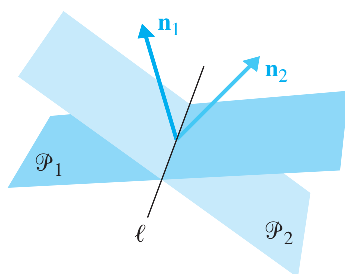


Figura 1.63
La intersección de dos planos es una recta

Comentarios

- Un plano es un objeto bidimensional, y su ecuación, en forma vectorial o paramétrica, requiere *dos* parámetros.
- Como muestra la figura 1.59, dado un punto P y un vector distinto de cero $\mathbf{n}$ en $\mathbb{R}^3$, existen infinitas rectas a través de P con $\mathbf{n}$ como vector normal. Sin embargo, P y dos vectores normales no paralelos $\mathbf{n}_1$ y $\mathbf{n}_2$ sirven para ubicar una recta ℓ de manera única, pues ℓ debe ser entonces la recta que pasa a través de P perpendicular al plano con ecuación $\mathbf{x} = \mathbf{p} + s\mathbf{n}_1 + t\mathbf{n}_2$ (figura 1.62). En consecuencia, una recta en $\mathbb{R}^3$ también puede especificarse mediante un par de ecuaciones

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \end{aligned}$$

que corresponden a cada vector normal. Pero, dado que dichas ecuaciones corresponden a un par de planos no paralelos (¿por qué no paralelos?), esta es justo la descripción de una recta como la intersección de dos planos no paralelos (figura 1.63). Algebraicamente, la recta consiste de todos los puntos (x, y, z) que satisfacen simultáneamente ambas ecuaciones. Este concepto se explorará más en el capítulo 2, cuando se estudie la solución de los sistemas de ecuaciones lineales.

Las tablas 1.2 y 1.3 resumen la información presentada hasta el momento acerca de las ecuaciones de rectas y planos.

Observe una vez más que una sola ecuación (general) describe una recta en $\mathbb{R}^2$, pero un plano en $\mathbb{R}^3$. [En dimensiones superiores, un objeto (recta, plano, etcétera) determinado por una sola ecuación de este tipo por lo general se conoce como ***hiperplano***.] La

Tabla 1.2 Ecuaciones de rectas en $\mathbb{R}^2$

Forma normal	Forma general	Forma vectorial	Forma paramétrica
$\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$	$ax + by = c$	$\mathbf{x} = \mathbf{p} + t\mathbf{d}$	$\begin{cases} x = p_1 + td_1 \\ y = p_2 + td_2 \end{cases}$

Tabla 1.3 Rectas y planos en $\mathbb{R}^3$

	Forma normal	Forma general	Forma vectorial	Forma paramétrica
Rectas	$\begin{cases} \mathbf{n}_1 \cdot \mathbf{x} = \mathbf{n}_1 \cdot \mathbf{p}_1 \\ \mathbf{n}_2 \cdot \mathbf{x} = \mathbf{n}_2 \cdot \mathbf{p}_2 \end{cases}$	$\begin{cases} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{cases}$	$\mathbf{x} = \mathbf{p} + t\mathbf{d}$	$\begin{cases} x = p_1 + td_1 \\ y = p_2 + td_2 \\ z = p_3 + td_3 \end{cases}$
Planos	$\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$	$ax + by + cz = d$	$\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$	$\begin{cases} x = p_1 + su_1 + tv_1 \\ y = p_2 + su_2 + tv_2 \\ z = p_3 + su_3 + tv_3 \end{cases}$

relación entre la dimensión del objeto, el número de ecuaciones requeridas y la dimensión del espacio está dada por la "fórmula de balanceo":

(*dimensión* del objeto) + (*número* de ecuaciones generales) = dimensión del espacio

Mientras más grande sea la dimensión del objeto, menos ecuaciones se necesitan. Por ejemplo, un plano en $\mathbb{R}^3$ es bidimensional, requiere una ecuación general y existe en un espacio tridimensional: $2 + 1 = 3$. Una recta en $\mathbb{R}^3$ es unidimensional y por tanto necesita $3 - 1 = 2$ ecuaciones. Note que la dimensión del objeto también concuerda con el número de parámetros en su forma vectorial o paramétrica. Las nociones de "dimensión" se aclararán en los capítulos 3 y 6, pero por el momento le servirán estas observaciones intuitivas.

Ahora puede encontrar la distancia desde un punto hasta una recta o un plano al combinar los resultados de la sección 1.2 con los de esta sección.

Ejemplo 1.32

Encuentre la distancia desde el punto $B = (1, 0, 2)$ hasta la recta ℓ a través del punto $A = (3, 1, 1)$ con vector director $\mathbf{d} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$.

Solución Como ya se determinó, es necesario calcular la longitud de $\overrightarrow{PB}$, donde P es el punto sobre ℓ al pie de la perpendicular desde B. Si se etiqueta $\mathbf{v} = \overrightarrow{AB}$, entonces $\overrightarrow{AP} = \text{proy}_{\mathbf{d}}(\mathbf{v})$ y $\overrightarrow{PB} = \mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v})$ (vea la figura 1.64). Los cálculos necesarios se realizaron en varios pasos.

Paso 1: $\mathbf{v} = \overrightarrow{AB} = \mathbf{b} - \mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$

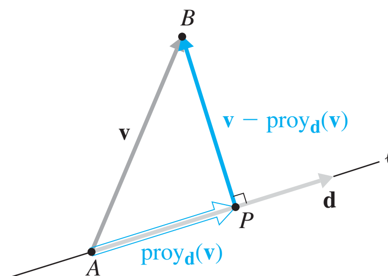


Figura 1.64
$d(B, \ell) = \|\mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v})\|$

Paso 2: La proyección de **v** sobre **d** es

$$\begin{aligned}\text{proy}_{\mathbf{d}}(\mathbf{v}) &= \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d}\\ &= \left(\frac{(-1)\cdot(-2) + 1\cdot(-1) + 0\cdot 1}{(-1)^2 + 1 + 0}\right)\begin{bmatrix}-1\\1\\0\end{bmatrix}\\ &= \tfrac{1}{2}\begin{bmatrix}-1\\1\\0\end{bmatrix}\\ &= \begin{bmatrix}-\frac{1}{2}\\ \frac{1}{2}\\0\end{bmatrix}\end{aligned}$$

Paso 3: El vector que se quiere es

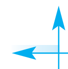
$$\mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v}) = \begin{bmatrix}-2\\-1\\1\end{bmatrix} - \begin{bmatrix}-\frac{1}{2}\\ \frac{1}{2}\\0\end{bmatrix} = \begin{bmatrix}-\frac{3}{2}\\ -\frac{3}{2}\\1\end{bmatrix}$$

Paso 4: La distancia $d(B, \ell)$ desde B hasta ℓ es

$$\|\mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v})\| = \left\|\begin{bmatrix}-\frac{3}{2}\\ -\frac{3}{2}\\1\end{bmatrix}\right\|$$

Al usar el Teorema 1.3(b) para simplificar el cálculo se obtiene

$$\begin{aligned}\|\mathbf{v} - \text{proy}_{\mathbf{d}}(\mathbf{v})\| &= \tfrac{1}{2}\left\|\begin{bmatrix}-3\\-3\\2\end{bmatrix}\right\|\\ &= \tfrac{1}{2}\sqrt{9 + 9 + 4}\\ &= \tfrac{1}{2}\sqrt{22}\end{aligned}$$



Nota

- En términos de nuestra notación anterior, $d(B, \ell) = d(\mathbf{v}, \text{proy}_{\mathbf{d}}(\mathbf{v}))$.

En el caso donde la recta ℓ está en $\mathbb{R}^2$ y su ecuación tiene la forma general $ax + by = c$, la distancia $d(B, \ell)$ desde $B = (x_0, y_0)$está dada por la fórmula

$$d(B, \ell) = \frac{|ax_0 + by_0 - c|}{\sqrt{a^2 + b^2}} \qquad (3)$$

Se le invita a demostrar esta fórmula en el ejercicio 39.

Ejemplo 1.33

Encuentre la distancia desde el punto $B = (1, 0, 2)$ hasta el plano $\mathcal{P}$ cuya ecuación general es $x + y - z = 1$.

Solución En este caso es necesario calcular la longitud de $\overrightarrow{PB}$, donde P es el punto sobre $\mathcal{P}$ al pie de la perpendicular desde B. Como muestra la figura 1.65, si A es cualquier punto sobre $\mathcal{P}$ y el vector normal $\mathbf{n} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ de $\mathcal{P}$ se sitúa de modo que su origen esté en A, entonces es necesario encontrar la longitud de la proyección de $\overrightarrow{AB}$ sobre $\mathbf{n}$. Nuevamente, los cálculos necesarios se hacen en pasos.

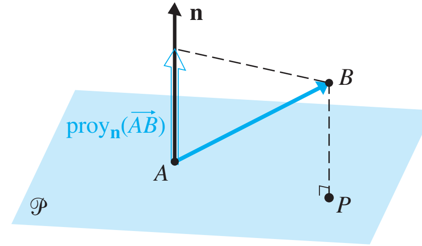


Figura 1.65
$d(B, \mathcal{P}) = \|\text{proy}_{\mathbf{n}}(\overrightarrow{AB})\|$

Paso 1: Mediante ensayo y error se encuentra cualquier punto cuyas coordenadas satisfagan la ecuación $x + y - z = 1$. $A = (1, 0, 0)$ lo hará.

Paso 2: Sea

$$\mathbf{v} = \overrightarrow{AB} = \mathbf{b} - \mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

Paso 3: La proyección de $\mathbf{v}$ sobre $\mathbf{n}$ es

$$\begin{aligned} \text{proy}_{\mathbf{n}}(\mathbf{v}) &= \left(\frac{\mathbf{n} \cdot \mathbf{v}}{\mathbf{n} \cdot \mathbf{n}}\right)\mathbf{n} \\ &= \left(\frac{1 \cdot 0 + 1 \cdot 0 - 1 \cdot 2}{1 + 1 + (-1)^2}\right)\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \\ &= -\tfrac{2}{3}\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -\frac{2}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \end{aligned}$$

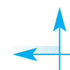
Paso 4: La distancia $(B, \mathcal{P})$desde B hasta $\mathcal{P}$ es

$$\begin{aligned} \|\text{proy}_{\mathbf{n}}(\mathbf{v})\| &= \left|-\tfrac{2}{3}\right| \left\| \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\| \\ &= \tfrac{2}{3} \left\| \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\| \\ &= \tfrac{2}{3}\sqrt{3} \end{aligned}$$

En general, la distancia $\mathrm{d}(B, \mathcal{P})$ desde el punto $B = (x_0, y_0, z_0)$ hasta el plano cuya ecuación general es $ax + by + cz = d$ está dada por la fórmula

$$\mathrm{d}(B, \mathcal{P}) = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}} \tag{4}$$

En el ejercicio 40 se le pedirá deducir esta fórmula.



Ejercicios 1.3

En los ejercicios 1 y 2, escriba la ecuación de la recta que pasa por P con vector normal **n** *en (a) forma normal y (b) forma general.*

1. $P = (0, 0)$, $\mathbf{n} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ **2.** $P = (2, 1)$, $\mathbf{n} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$

En los ejercicios 3-6, escriba la ecuación de la recta que pasa por P con vector director **d** *en (a) forma vectorial y (b) forma paramétrica.*

3. $P = (1, 0)$, $\mathbf{d} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ **4.** $P = (3, -3)$, $\mathbf{d} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

5. $P = (0, 0, 0)$, $\mathbf{d} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ **6.** $P = (-3, 1, 2)$, $\mathbf{d} = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$

En los ejercicios 7 y 8, escriba la ecuación del plano que pasa por P con vector normal **n** *en (a) forma normal y (b) forma general.*

7. $P = (0, 1, 0)$, $\mathbf{n} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$ **8.** $P = (-3, 1, 2)$, $\mathbf{n} = \begin{bmatrix} 1 \\ 0 \\ 5 \end{bmatrix}$

En los ejercicios 9 y 10, escriba la ecuación del plano que pasa por P con vectores directores **u** *y* **v** *en (a) forma vectorial y (b) forma paramétrica.*

9. $P = (0, 0, 0)$, $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$

10. $P = (4, -1, 3)$, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 11 y 12, proporcione la ecuación vectorial de la recta que pasa por P y Q.

11. $P = (1, -2)$, $Q = (3, 0)$

12. $P = (4, -1, 3)$, $Q = (2, 1, 3)$

En los ejercicios 13 y 14, proporcione la ecuación vectorial del plano que pasa por P, Q y R.

13. $P = (1, 1, 1)$, $Q = (4, 0, 2)$, $R = (0, 1, -1)$

14. $P = (1, 0, 0)$, $Q = (0, 1, 0)$, $R = (0, 0, 1)$

15. Encuentre las ecuaciones paramétricas y una ecuación en forma vectorial para las rectas en $\mathbb{R}^2$ con las siguientes ecuaciones:

(a) $y = 3x - 1$ **(b)** $3x + 2y = 5$

16. Considere la ecuación vectorial $\mathbf{x} = \mathbf{p} + t(\mathbf{q} - \mathbf{p})$, donde $\mathbf{p}$ y $\mathbf{q}$ corresponden a distintos puntos P y Q en $\mathbb{R}^2$ o $\mathbb{R}^3$.

(a) Demuestre que esta ecuación describe el segmento de recta $\overline{PQ}$ cuando t varía de 0 a 1.

(b) ¿Para cuál valor de t, $\mathbf{x}$ es el punto medio de $\overline{PQ}$ y cuál es $\mathbf{x}$ en este caso?

(c) Encuentre el punto medio de $\overline{PQ}$ cuando $P = (2, -3)$ y $Q = (0, 1)$.

(d) Encuentre el punto medio de $\overline{PQ}$ cuando $P = (1, 0, 1)$ y $Q = (4, 1, -2)$.

(e) Encuentre los dos puntos que dividen $\overline{PQ}$ del inciso (c) en tres partes iguales.

(f) Encuentre los dos puntos que dividen $\overline{PQ}$ del inciso (d) en tres partes iguales.

17. Sugiera una "demostración vectorial" del hecho de que, en $\mathbb{R}^2$, dos rectas con pendientes m_1 y m_2 son perpendiculares si y sólo si $m_1 m_2 = -1$.

18. La recta ℓ pasa a través del punto $P = (1, -1, 1)$ y tiene vector director $\mathbf{d} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$. Para cada uno de los siguientes planos $\mathcal{P}$, determine si ℓ y $\mathcal{P}$ son paralelos, perpendiculares o ninguno de los dos.

(a) $2x + 3y - z = 1$ **(b)** $4x - y + 5z = 0$
(c) $x - y - z = 3$ **(d)** $4x + 6y - 2z = 0$

19. El plano $\mathcal{P}_1$ tiene la ecuación $4x - y + 5z = 2$. Para cada uno de los planos $\mathcal{P}$ del ejercicio 18, determine si $\mathcal{P}_1$ y $\mathcal{P}$ son paralelos, perpendiculares o ninguno de los dos.

20. Encuentre la forma vectorial de la ecuación de la recta en $\mathbb{R}^2$ que pasa a través de $P = (2, -1)$ y es perpendicular a la recta con ecuación general $2x - 3y = 1$.

21. Encuentre la forma vectorial de la ecuación de la recta en $\mathbb{R}^2$ que pasa a través de $P = (2, -1)$ y es paralela a la recta con ecuación general $2x - 3y = 1$.

22. Encuentre la forma vectorial de la ecuación de la recta en $\mathbb{R}^3$ que pasa a través de $P = (-1, 0, 3)$ y es perpendicular al plano con ecuación general $x - 3y + 2z = 5$.

23. Encuentre la forma vectorial de la ecuación de la recta en $\mathbb{R}^3$ que pasa a través de $P = (-1, 0, 3)$ y es paralela a la recta con ecuaciones paramétricas

$$\begin{aligned} x &= 1 - t \\ y &= 2 + 3t \\ z &= -2 - t \end{aligned}$$

24. Encuentre la forma normal de la ecuación del plano que pasa por $P = (0, -2, 5)$ y es paralelo al plano con ecuación general $6x - y + 2z = 3$.

25. Un cubo tiene vértices en los ocho puntos (x, y, z), donde x, y y z son 0 o 1. (Vea la figura 1.34.)

(a) Encuentre las ecuaciones generales de los planos que determinan las seis caras (lados) del cubo.

(b) Encuentre la ecuación general del plano que con tiene la diagonal desde el origen hasta(1, 1, 1) y es perpendicular al plano xy.

(c) Encuentre la ecuación general del plano que contiene las diagonales laterales a las que se refiere el ejemplo 1.22.

26. Encuentre la ecuación del conjunto de todos los puntos que son equidistantes de los puntos $P = (1, 0, -2)$ y $Q = (5, 2, 4)$.

En los ejercicios 27 y 28, encuentre la distancia desde el punto Q hasta la recta ℓ.

27. $Q = (2, 2)$, ℓ con ecuación $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} + t\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

28. $Q = (0, 1, 0)$, ℓ con ecuación $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} -2 \\ 0 \\ 3 \end{bmatrix}$

En los ejercicios 29 y 30, encuentre la distancia desde el punto Q hasta el plano $\mathcal{P}$.

29. $Q = (2, 2, 2)$, $\mathcal{P}$ con ecuación $x + y - z = 0$

30. $Q = (0, 0, 0)$, $\mathcal{P}$ con ecuación $x - 2y + 2z = 1$

La figura 1.66 sugiere una forma de usar vectores para localizar el punto R sobre ℓ que está más cerca de Q.

31. Encuentre el punto R sobre ℓ que esté más cerca de Q en el ejercicio 27.

32. Encuentre el punto R sobre ℓ que esté más cerca de Q en el ejercicio 28.

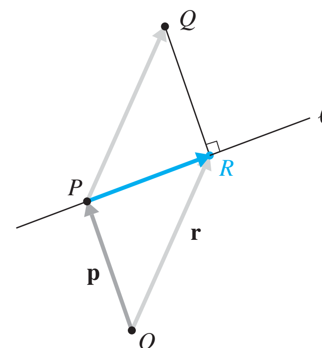


Figura 1.66
$\mathbf{r} = \mathbf{p} + \overrightarrow{PR}$

La figura 1.67 sugiere una forma de usar vectores para localizar el punto R sobre $\mathcal{P}$ que esté más cerca de Q.

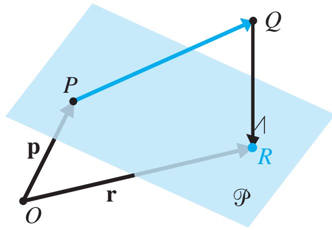


Figura 1.67

$\mathbf{r} = \mathbf{p} + \overrightarrow{PQ} + \overrightarrow{QR}$

33. Encuentre el punto R sobre $\mathcal{P}$ que esté más cerca de Q en el ejercicio 29.

34. Encuentre el punto R sobre $\mathcal{P}$ que esté más cerca de Q en el ejercicio 30.

En los ejercicios 35 y 36, encuentre la distancia entre las rectas paralelas.

35. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} -2 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix} + t\begin{bmatrix} -2 \\ 3 \end{bmatrix}$

36. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 37 y 38, encuentre la distancia entre los planos paralelos.

37. $2x + y - 2z = 0$ y $2x + y - 2z = 5$

38. $x + y + z = 1$ y $x + y + z = 3$

39. Demuestre la ecuación (3) de la página 43.

40. Demuestre la ecuación (4) de la página 44.

41. Demuestre que, en $\mathbb{R}^2$, la distancia entre rectas paralelas con ecuaciones $\mathbf{n} \cdot \mathbf{x} = c_1$ y $\mathbf{n} \cdot \mathbf{x} = c_2$ está dada por

$$\frac{|c_1 - c_2|}{\|\mathbf{n}\|}.$$

42. Demuestre que la distancia entre planos paralelos con ecuaciones $\mathbf{n} \cdot \mathbf{x} = d_1$ y $\mathbf{n} \cdot \mathbf{x} = d_2$ está dada por

$$\frac{|d_1 - d_2|}{\|\mathbf{n}\|}.$$

Si dos planos no paralelos $\mathcal{P}_1$ y $\mathcal{P}_2$ tienen vectores normales $\mathbf{n}_1$ y $\mathbf{n}_2$ y θ es el ángulo entre $\mathbf{n}_1$ y $\mathbf{n}_2$, entonces se define el ángulo entre $\mathcal{P}_1$ y $\mathcal{P}_2$ como θ o como $180° - \theta$, cualquiera que sea un ángulo agudo. (Figura 1.68)

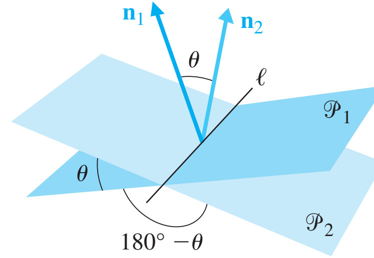


Figura 1.68

En los ejercicios 43-44, encuentre el ángulo agudo entre los planos con las ecuaciones dadas.

43. $x + y + z = 0$ y $2x + y - 2z = 0$

44. $3x - y + 2z = 5$ y $x + 4y - z = 2$

En los ejercicios 45-46, demuestre que el plano y la recta con las ecuaciones dadas se intersectan, y luego encuentre el ángulo agudo de intersección entre ellos.

45. El plano dado por $x + y + 2z = 0$ y la recta dada por

$$\begin{aligned} x &= 2 + t \\ y &= 1 - 2t \\ z &= 3 + t \end{aligned}$$

46. El plano dado por $4x - y - z = 6$ y la recta dada por

$$\begin{aligned} x &= t \\ y &= 1 + 2t \\ z &= 2 + 3t \end{aligned}$$

Los ejercicios 47-48 exploran un planteamiento para el problema de encontrar la proyección de un vector sobre un plano. Como muestra la figura 1.69, si $\mathcal{P}$ es un plano a través del origen en $\mathbb{R}^3$ con vector normal $\mathbf{n}$, y $\mathbf{v}$ es un vector en $\mathbb{R}^3$, entonces $\mathbf{p} = \text{proy}_{\mathcal{P}}(\mathbf{v})$ es un vector en $\mathcal{P}$ tal que $\mathbf{v} - c\mathbf{n} = \mathbf{p}$ para algún escalar c.

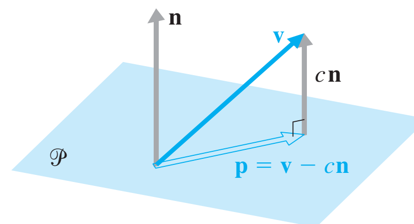


Figura 1.69
Proyección sobre un plano

47. Con el hecho de que **n** es ortogonal a todo vector en $\mathscr{P}$ (y por tanto a **p**), calcule c y en consecuencia encuentre una expresión para **p** en términos de **v** y **n**.

48. Use el método del ejercicio 43 para encontrar la proyección de

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$$

sobre los planos con las siguientes ecuaciones:

(a) $x + y + z = 0$ **(b)** $3x - y + z = 0$

(c) $x - 2z = 0$ **(d)** $2x - 3y + z = 0$

Exploración

El producto cruz

Sería conveniente que se pudiera convertir fácilmente la forma vectorial $\mathbf{x} = \mathbf{p} + s\mathbf{u} + t\mathbf{v}$ de la ecuación de un plano a la forma normal $\mathbf{n} \cdot \mathbf{x} = \mathbf{n} \cdot \mathbf{p}$. Lo que se necesita es un proceso que, dados dos vectores no paralelos $\mathbf{u}$ y $\mathbf{v}$, produzca un tercer vector $\mathbf{n}$ que sea ortogonal tanto a $\mathbf{u}$ como a $\mathbf{v}$. Un método es usar una construcción conocida como ***producto cruz*** de vectores. *Sólo válido en* $\mathbb{R}^3$, se define del modo siguiente:

Definición El ***producto cruz*** de $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ es el vector $\mathbf{u} \times \mathbf{v}$ definido por

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$$

A continuación se ilustra un atajo que puede ayudarle a recordar cómo calcular el producto cruz de dos vectores. Abajo de cada vector completo escriba sus primeros dos componentes. Ignore los dos componentes en la línea superior y considere cada bloque de cuatro: reste los productos de los componentes conectados con líneas discontinuas, de los productos de los componentes conectados con líneas sólidas. (Es útil notar que el primer componente de $\mathbf{u} \times \mathbf{v}$ no tiene números 1 como subíndices, el segundo no tiene 2 y el tercero no tiene 3.)

$$\begin{array}{lll} u_1 & v_1 & \\ u_2 & v_2 & \\ u_3 & v_3 & u_2v_3 - u_3v_2 \\ u_1 & v_1 & u_3v_1 - u_1v_3 \\ u_2 & v_2 & u_1v_2 - u_2v_1 \end{array}$$

Los siguientes problemas exploran brevemente el producto cruz.

1. Calcule $\mathbf{u} \times \mathbf{v}$.

(a) $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$ (b) $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(c) $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -4 \\ -6 \end{bmatrix}$ (d) $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

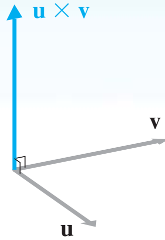


Figura 1.70

2. Demuestre que $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$, y $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$.

3. Con la definición de producto cruz, demuestre que $\mathbf{u} \times \mathbf{v}$ (como se muestra en la figura 1.70) es ortogonal a $\mathbf{u}$ y $\mathbf{v}$.

4. Use el producto cruz como ayuda para encontrar la forma normal de la ecuación del plano.

(a) El plano que pasa por $P = (1, 0, -2)$, paralelo a $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$

(b) El plano que pasa por $P = (0, -1, 1)$, $Q = (2, 0, 2)$ y $R = (1, 2, -1)$

5. Demuestre las siguientes propiedades del producto cruz:

(a) $\mathbf{v} \times \mathbf{u} = -(\mathbf{u} \times \mathbf{v})$ (b) $\mathbf{u} \times \mathbf{0} = \mathbf{0}$
(c) $\mathbf{u} \times \mathbf{u} = \mathbf{0}$ (d) $\mathbf{u} \times k\mathbf{v} = k(\mathbf{u} \times \mathbf{v})$
(e) $\mathbf{u} \times k\mathbf{u} = \mathbf{0}$ (f) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$

6. Demuestre las siguientes propiedades del producto cruz:

(a) $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$ (b) $\mathbf{u} \times (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \cdot \mathbf{w})\mathbf{v} - (\mathbf{u} \cdot \mathbf{v})\mathbf{w}$
(c) $\|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2\|\mathbf{v}\|^2 - (\mathbf{u} \cdot \mathbf{v})^2$

7. Vuelva a trabajar los problemas 2 y 3, esta vez usando los problemas 5 y 6.

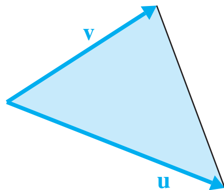


Figura 1.71

8. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en $\mathbb{R}^3$ y sea θ el ángulo entre $\mathbf{u}$ y $\mathbf{v}$.

(a) Demuestre que $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\|\|\mathbf{v}\| \operatorname{sen} \theta$. [*Sugerencia:* use el problema 6(c).]

(b) Demuestre que el área $\mathcal{A}$ del triángulo determinado por $\mathbf{u}$ y $\mathbf{v}$ (como se muestra en la figura 1.71) está dada por

$$\mathcal{A} = \tfrac{1}{2}\|\mathbf{u} \times \mathbf{v}\|$$

(c) Use el resultado del inciso (b) para calcular el área del triángulo con vértices $A = (1, 2, 1)$, $B = (2, 1, 0)$ y $C = (5, -1, 3)$.

1.4 Aplicaciones

Vectores fuerza

Los vectores pueden usarse para modelar fuerzas. Por ejemplo, un viento que sopla a 30 km/h en dirección hacia el oeste o la gravedad de la Tierra que actúa sobre una masa de 1 kg con una fuerza de 9.8 newtons hacia abajo se representan mejor mediante vectores, pues cada uno consiste de una magnitud y una dirección.

La fuerza se define como el producto de masa por aceleración debida a la gravedad (que, en la Tierra, es de 9.8 m/s^2). Por tanto, una masa de 1 kg ejerce una fuerza descendente de 1 kg $\times$ 9.8 m/s^2 o 9.8 kg·m/s^2. Esta unidad de medición es un ***newton*** (N). De este modo, la fuerza ejercida por una masa de 1 kg es 9.8 N.

Con frecuencia, múltiples fuerzas actúan sobre un objeto. En tales situaciones, el resultado neto de todas las fuerzas que actúan en conjunto es una sola fuerza llamada ***resultante***, que es simplemente la suma vectorial de las fuerzas individuales (figura 1.72). Cuando varias fuerzas actúan sobre un objeto, es posible que la fuerza resultante sea cero. En este caso, el objeto claramente no se mueve en alguna dirección y se dice que está en ***equilibrio***. Cuando un objeto está en equilibrio y los vectores fuerza que actúan sobre él se ordenan de origen a cola, el resultado es un polígono cerrado (figura 1.73).

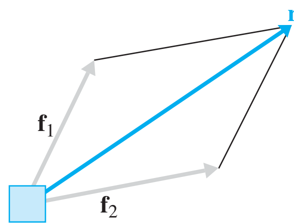


Figura 1.72
La resultante de dos fuerzas

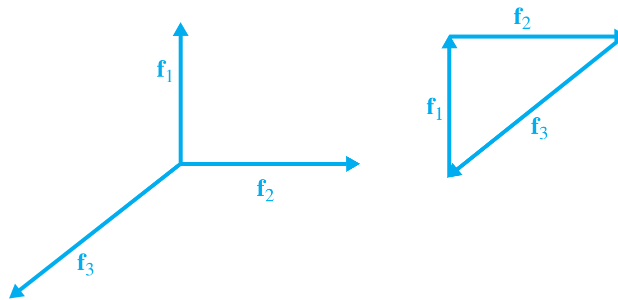


Figura 1.73
Equilibrio

Ejemplo 1.34

Ann y Bert tratan de rodar una piedra por el camino. Ann empuja con una fuerza de 20 N en dirección hacia el norte, mientras que Bert empuja con una fuerza de 40 N en dirección al este.

(a) ¿Cuál es la fuerza resultante sobre la piedra?
(b) Carla trata de evitar que Ann y Bert muevan la piedra. ¿Qué fuerza debe aplicar Carla para mantener la piedra en equilibrio?

Solución (a) La figura 1.74 muestra la posición de las dos fuerzas. Con la regla del paralelogramo, sume las dos fuerzas para obtener la resultante **r** como se muestra. Por el teo-

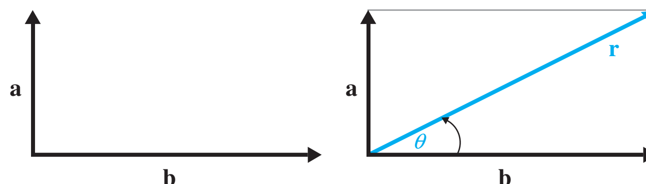


Figura 1.74
La resultante de dos fuerzas

rema de Pitágoras, se ve que $\|\mathbf{r}\| = \sqrt{20^2 + 40^2} = \sqrt{2000} \approx 44.72$ N. Para la dirección de **r**, calcule el ángulo θ entre **r** y la fuerza hacia el este de Bert. Se encuentra que sen $\theta = 20/\|\mathbf{r}\| \approx 0.447$, de modo que $\theta \approx 26.57°$

(b) Si denota las fuerzas ejercidas por Ann, Bert y Carla como **a**, **b** y **c**, respectivamente, entonces se requiere que $\mathbf{a} + \mathbf{b} + \mathbf{c} = \mathbf{0}$. Por tanto, $\mathbf{c} = -(\mathbf{a} + \mathbf{b}) = -\mathbf{r}$, de modo que Carla debe ejercer una fuerza de 44.72 N en la dirección opuesta a **r**.

Con frecuencia, se está interesado en la descomposición de un vector fuerza en otros vectores cuya resultante es el vector dado. Este proceso se conoce como ***descomposición de un vector en componentes***. En dos dimensiones, se quiere descomponer un vector en dos componentes. Sin embargo, existen infinitas formas de hacerlo; la más útil será descomponer el vector en dos componentes *ortogonales*. (Los capítulos 5 y 7 exploran esta idea con más generalidad.) Por lo general, esto se hace al introducir ejes coordenados y elegir los componentes de modo que uno sea paralelo al eje x y el otro al eje y. Comúnmente, a dichos componentes se les conoce como componentes horizontal y vertical, respectivamente. En la figura 1.75, **f** es el vector dado y $\mathbf{f}_x$ y $\mathbf{f}_y$ son sus componentes horizontal y vertical.

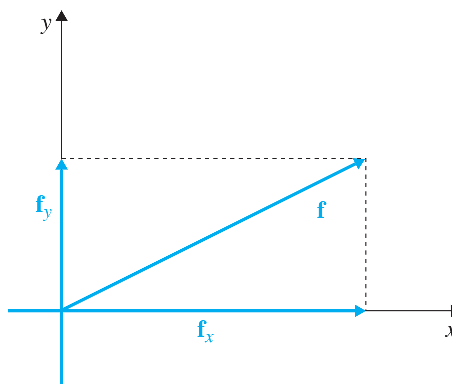


Figura 1.75
Descomposición de un vector en componentes

Ejemplo 1.35

Ann jala el manubrio de un carro con una fuerza de 100 N. Si el manubrio forma un ángulo de 20° con la horizontal, ¿cuál es la fuerza que tiende a jalar al carro hacia adelante y cuál fuerza tiende a levantarlo del suelo?

Solución La figura 1.76 muestra la situación y el diagrama vectorial que necesita considerar.

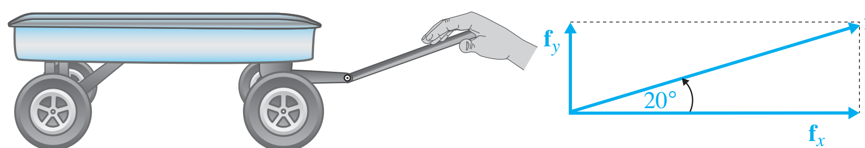
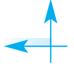


Figura 1.76

Se ve que

$$\|\mathbf{f}_x\| = \|\mathbf{f}\| \cos 20^\circ \text{ y } \|\mathbf{f}_y\| = \|\mathbf{f}\| \operatorname{sen} 20^\circ$$

Por tanto, $\|\mathbf{f}_x\| \approx 100(0.9397) \approx 93.97$ y $\|\mathbf{f}_y\| \approx 100(0.3420) \approx 34.20$. De modo que el carro se jala hacia adelante con una fuerza de aproximadamente 93.97 N y tiende a elevarse del suelo con una fuerza de aproximadamente 34.20 N.



El siguiente ejemplo se resuelve usando dos métodos diferentes. La primera solución considera un triángulo de fuerzas en equilibrio; la segunda solución usa descomposición de fuerzas en componentes.

Ejemplo 1.36

La figura 1.77 muestra una pintura que cuelga del techo mediante dos alambres. Si la pintura tiene una masa de 5 kg y si los dos cables forman ángulos de 45 y 60 grados con el techo, determine la tensión en cada cable.

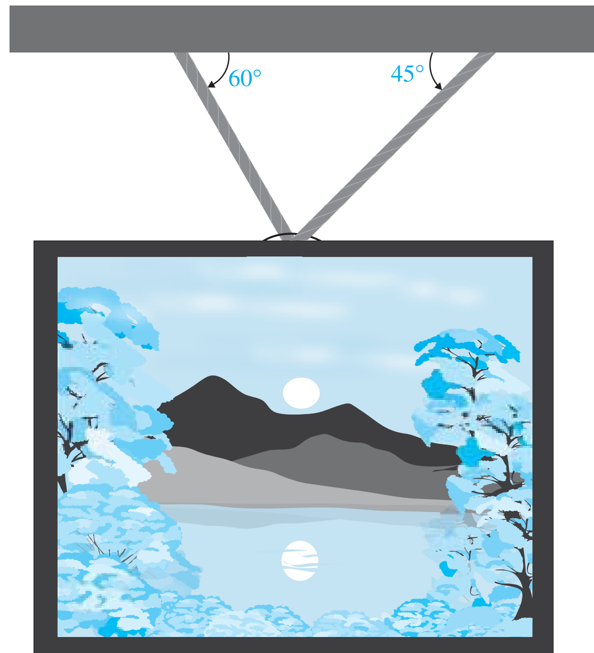


Figura 1.77

Solución 1 Suponga que la pintura está en equilibrio. entonces los dos cables deben proporcionar suficiente fuerza ascendente para equilibrar la fuerza descendente de la gravedad. La gravedad ejerce una fuerza descendente de $5 \times 9.8 = 49$ N sobre la pintura, de modo que los dos cables deben jalar conjuntamente hacia arriba con 49 N de fuerza. Sean $\mathbf{f}_1$ y $\mathbf{f}_2$ las tensiones en los cables y sea $\mathbf{r}$ su resultante (figura 1.78). Se sigue que $\|\mathrm{r}\| = 49$, pues está en equilibrio.

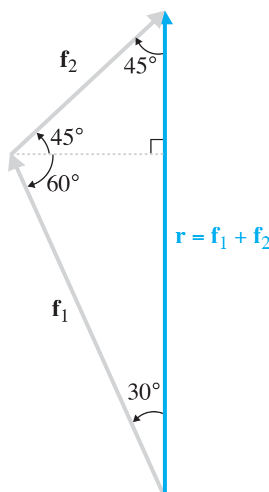


Figura 1.78

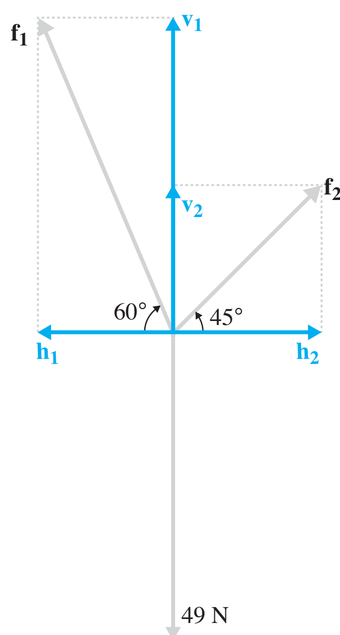


Figura 1.79

Al usar la ley de los senos, se tiene

$$\frac{\|\mathbf{f}_1\|}{\text{sen } 45°} = \frac{\|\mathbf{f}_2\|}{\text{sen } 30°} = \frac{\|\mathbf{r}\|}{\text{sen } 105°}$$

de modo que

$$\|\mathbf{f}_1\| = \frac{\|\mathbf{r}\| \text{ sen } 45°}{\text{sen } 105°} \approx \frac{49(0.7071)}{0.9659} \approx 35.87 \text{ y } \|\mathbf{f}_2\| = \frac{\|\mathbf{r}\| \text{ sen } 30°}{\text{sen } 105°} \approx \frac{49(0.5)}{0.9659} \approx 25.36$$

Por tanto, las tensiones en los cables son aproximadamente 35.8 N y 25.36 N.

Solución 2 Descomponga $\mathbf{f}_1$ y $\mathbf{f}_2$ en componentes horizontales y verticales, por decir, $\mathbf{f}_1 = \mathbf{h}_1 + \mathbf{v}_1$ y $\mathbf{f}_2 = \mathbf{h}_2 + \mathbf{v}_2$, y note que, como arriba, hay una fuerza descendente de 49 N (figura 1.79).

Se sigue que

$$\|\mathbf{h}_1\| = \|\mathbf{f}_1\| \cos 60° = \frac{\|\mathbf{f}_1\|}{2}, \quad \|\mathbf{v}_1\| = \|\mathbf{f}_1\| \text{ sen } 60° = \frac{\sqrt{3}\|\mathbf{f}_1\|}{2},$$

$$\|\mathbf{h}_2\| = \|\mathbf{f}_2\| \cos 45° = \frac{\|\mathbf{f}_2\|}{\sqrt{2}}, \quad \|\mathbf{v}_2\| = \|\mathbf{f}_2\| \text{ sen } 45° = \frac{\|\mathbf{f}_2\|}{\sqrt{2}}$$

Dado que la pintura está en equilibrio, los componentes horizontales deben equilibrarse, al igual que los componentes verticales. Por tanto, $\|\mathbf{h}_1\| = \|\mathbf{h}_2\|$ y $\|\mathbf{v}_1\| + \|\mathbf{v}_2\| = 49$, de donde se concluye que

$$\|\mathbf{f}_1\| = \frac{2\|\mathbf{f}_2\|}{\sqrt{2}} = \sqrt{2}\|\mathbf{f}_2\| \qquad \text{y} \qquad \frac{\sqrt{3}\|\mathbf{f}_1\|}{2} + \frac{\|\mathbf{f}_2\|}{\sqrt{2}} = 49$$

Al sustituir la primera de estas ecuaciones en la segunda ecuación se obtiene

$$\frac{\sqrt{3}\|\mathbf{f}_2\|}{\sqrt{2}} + \frac{\|\mathbf{f}_2\|}{\sqrt{2}} = 49, \quad \text{o} \quad \|\mathbf{f}_2\| = \frac{49\sqrt{2}}{1 + \sqrt{3}} \approx 25.36$$

Por tanto, $\|\mathbf{f}_1\| = \sqrt{2}\|\mathbf{f}_2\| \approx 1.4142(25.36) \approx 35.87$, de modo que las tensiones en los cables son aproximadamente 35.87 N y 25.36 N, como antes.

Vectores código

A lo largo de la historia, la gente ha transmitido información usando códigos. En ocasiones la intención es disfrazar el mensaje a enviar, como cuando cada letra en una palabra se sustituye con otra diferente de acuerdo con una regla de sustitución. Aunque fascinantes, dichos códigos secretos, o cifrados, no se tratarán aquí; pues son el objetivo del campo de la *criptografía*. En vez de ello, el texto se concentrará en los códigos que se usan cuando deben transmitirse datos de manera electrónica.

Un ejemplo familiar de tal código es el código Morse, con su sistema de puntos y rayas. La llegada de las computadoras digitales en el siglo XX condujo a la necesidad de transmitir grandes cantidades de datos rápidamente y con exactitud. Las computadoras se diseñaron para codificar datos como secuencias de números 0 y 1. Muchos avances tecnológicos recientes dependen de códigos y se les encuentra todos los días sin estar al tanto de ellos: comunicaciones satelitales, reproductores de discos compactos, los códi-

gos universales de producto (UPC) asociados con los códigos de barras que se encuentran en las mercancías y los números internacionales normalizados de los libros (ISBN) que se encuentran en todo libro publicado en la actualidad, son sólo algunos ejemplos.

En esta sección se usarán vectores para diseñar códigos que detecten errores que pueden ocurrir en la transmisión de datos. En capítulos posteriores se construirán códigos que pueden no sólo detectar, sino también corregir errores. Los vectores que surgen en el estudio de los códigos no son vectores en $\mathbb{R}^n$, sino en $\mathbb{Z}_2^n$ o, de manera más general, $\mathbb{Z}_m^n$. La primera vez que se encontraron tales vectores fue en la sección 1.1. Dado que las computadoras representan los datos en términos de 0 y 1 (que pueden interpretarse como encendido/apagado, cerrado/abierto, falso/verdadero o no/sí), comience por considerar códigos binarios, que consisten de vectores con entradas en $\mathbb{Z}_2$.

En la práctica, se tiene un mensaje (que consiste de palabras, números o símbolos) que se quiere transmitir. Comience por codificar cada "palabra" del mensaje como un vector binario.

La moderna teoría de códigos se originó con el trabajo del matemático y científico de la computación, el estadounidense Claude Shannon (1916–2001), cuya tesis de 1937 mostró cómo el álgebra podía desempeñar un papel en el diseño y análisis de los circuitos eléctricos. Tiempo después, Shannon sería útil en la formación del campo de la *teoría de la información* y propondría las bases teóricas de lo que ahora se conoce como códigos de corrección de error.



Definición Un ***código binario*** es un conjunto de vectores binarios (de la misma longitud) llamados ***vectores código***. El proceso de convertir un mensaje en vectores código se llama ***codificación***, y el proceso inverso se llama ***decodificación***.

Como verá, es enormemente deseable que un código también tenga otras propiedades, como la capacidad de detenerse cuando ocurra un error en la transmisión de un vector código y, si es posible, sugerir cómo corregir el error.

Suponga que ya se codificó un mensaje como un conjunto de vectores código binario. Ahora se quiere enviar los vectores código binario a través de un *canal* (como un transmisor de radio, una línea telefónica, un cable de fibra óptica o un CD láser). Por desgracia, el canal puede ser "ruidoso" (debido a interferencia eléctrica, señales competidoras, o polvo y rayones). Como resultado, pueden introducirse errores: algunos de los 0 pueden cambiar por 1 y viceversa. ¿Cómo puede protegerse contra este problema?

Ejemplo 1.37

Se quiere codificar y transmitir un mensaje que consiste de una de las palabras *arriba*, *abajo*, *izquierda* o *derecha*. Se decide usar los cuatro vectores en $\mathbb{Z}_2^2$ como el código binario, como se muestra en la tabla 1.4.

Si el receptor también tiene esta tabla y el mensaje codificado se transmite sin error, la decodificación es trivial. Sin embargo, suponga que ocurrió un solo error. (Por error se entiende que cambió un componente del vector código.) Por ejemplo, suponga que se envía el mensaje "abajo" codificado como [0, 1], pero ocurre un error en la transmisión del primer componente y el 0 cambió por 1. Entonces el receptor vería [1, 1] en su lugar y decodificaría el mensaje como "derecha". (Sólo debe preocuparse por el caso de errores individuales como éste. En la práctica, por lo general se supone que la probabilidad de múltiples errores es despreciablemente pequeña.) Incluso si el receptor supiera (de alguna manera) que ocurrió un solo error, no sabría si el vector código correcto era [0, 1] o [1, 0].

Tabla 1.4

Mensaje	arriba	abajo	izquierda	derecha
Código	[0, 0]	[0, 1]	[1, 0]	[1, 1]

Pero suponga que se envía el mensaje usando un código que fuese un subconjunto de $\mathbb{Z}_2^3$; en otras palabras, un código binario de longitud 3, como se muestra en la tabla 1.5

Tabla 1.5

Mensaje	arriba	abajo	izquierda	derecha
Código	[0, 0, 0]	[0, 1, 1]	[1, 0, 1]	[1, 1, 0]

Este código puede detectar cualquier error individual. Por ejemplo, si se envió "abajo" como [1, 1, 1] y ocurrió un error en un componente, el receptor leería [1, 1, 1], [0, 0, 1] o [0, 1, 0], ninguno de los cuales es un vector de código. De modo que el receptor sabría que ocurrió un error (mas no dónde) y podría pedir la retransmisión del mensaje codificado. (¿Por qué el receptor no sabría dónde estuvo el error?)

El término *paridad* proviene de la palabra latina *par*, que significa "igual". Se dice que dos enteros tienen la misma paridad si ambos son pares o ambos son impares.

El código en la tabla 1.5 es un ejemplo de un ***código de detección de error***. Hasta la década de 1940, esto era lo mejor que podía lograrse. La llegada de las computadoras digitales condujo al desarrollo de códigos que podían *corregir*, así como detectar errores. Éstos se considerarán en los capítulos 3, 6 y 7.

El mismo mensaje a transmitir puede consistir de vectores binarios. En este caso, un simple pero útil código de detección de errores es un ***código de control de paridad***, que se crea al agregar un componente adicional, llamado ***dígito de control***, a cada vector, de modo que la paridad (el número total de números 1) sea par.

Ejemplo 1.38

Si el mensaje por enviar es el vector binario [1, 0, 0, 1, 0, 1], que tiene una cantidad impar de números 1, entonces el dígito de control será 1 (para hacer que el número total de 1 en el vector código sea par) y el vector código será [1, 0, 0, 1, 0, 1, 1].

Note que se detectará un solo error, pues ello causará que la paridad del vector código cambie de par a impar. Por ejemplo, si ocurriera un error en el tercer componente, el vector código se recibiría como [1, 0, 1, 1, 0, 1, 1], cuya paridad es impar porque tiene cinco números 1.

Observe este concepto de manera un poco más formal. Suponga que el mensaje es el vector binario $\mathbf{b} = [b_1, b_2, \ldots, b_n]$ en $\mathbb{Z}_2^n$. Entonces el vector código de control de paridad es $\mathbf{v} = [b_1, b_2, \ldots, b_n, d]$ en $\mathbb{Z}_2^{n+1}$, donde el dígito de control d se elige de modo que

$$b_1 + b_2 + \cdots + b_n + d = 0 \text{ en } \mathbb{Z}_2$$

o, de manera equivalente, de modo que

$$\mathbf{1} \cdot \mathbf{v} = 0$$

donde $\mathbf{1} = [1, 1, \ldots, 1]$, un vector cuyos componentes son todos los números 1. El vector $\mathbf{1}$ se llama ***vector de control***. Si el vector $\mathbf{v}'$ se recibe y $\mathbf{1} \cdot \mathbf{v}' = 1$, entonces puede estar seguro de que ocurrió un error. (Aunque no se considera la posibilidad de más de un error, observe que este esquema no detectará un número par de errores.)

Los códigos de control de paridad son un caso especial de los más generales ***códigos de dígito de control***, que se considerarán después de extender primero las ideas precedentes a

escenarios más generales. Los códigos que usan vectores *m*-arios se llaman ***códigos m-arios.*** El siguiente ejemplo es una extensión directa del ejemplo 1.39 a códigos ternarios.

Ejemplo 1.39

Sea $\mathbf{b} = [b_1, b_2, \ldots, b_n]$ un vector en $\mathbb{Z}_3^n$. Entonces, un vector código de dígito de control puede definirse mediante $\mathbf{v} = [b_1, b_2, \ldots, b_n, d]$ (en $\mathbb{Z}_3^{n+1}$),con el dígito de control d elegido de modo que

$$\mathbf{1} \cdot \mathbf{v} = 0$$

(donde el vector de control $\mathbf{1} = [1, 1, \ldots, 1]$ es el vector de números 1 en $\mathbb{Z}_3^{n+1}$); esto es, el dígito de control satisface

$$b_1 + b_2 + \cdots + b_n + d = 0 \text{ en } \mathbb{Z}_3$$

Por ejemplo, considere el vector $\mathbf{u} = [2, 2, 0, 1, 2]$. La suma de sus componentes es $2 + 2 + 0 + 1 + 2 = 1$, de modo que el dígito de control debe ser 2 (pues $1 + 2 = 0$). En consecuencia, el vector código asociado es $\mathbf{v} = [2, 2, 0, 1, 2, 2]$.

Aunque los códigos de dígito de control simples detectarán errores individuales, con frecuencia también es importante detectar otros tipos comunes de errores, como el intercambio accidental, o *transposición*, de dos componentes adyacentes. (Por ejemplo, la transposición del segundo y tercero componentes de **v** en el ejemplo 1.39 resultaría en el vector incorrecto $\mathbf{v}' = [2, 0, 2, 1, 2, 2]$.) Para tales propósitos, se diseñaron otros tipos de códigos de dígitos de control. Muchos de ellos simplemente sustituyen el vector de control **1** por algún otro vector **c** cuidadosamente elegido.

Ejemplo 1.40

El código universal de producto, o UPC (figura 1.80), es un código asociado con los códigos de barra que se encuentran en muchos tipos de mercancía.

Las barras negras y blancas que se escanean con un láser en un mostrador de verificación en una tienda, corresponden a un vector 10-ario $\mathbf{u} = [u_1, u_2, \ldots, u_{11}, d]$ de longitud 12. Los primeros 11 componentes forman un vector en $\mathbb{Z}_{10}^{11}$ que dan la información del fabricante y el producto; el último componente d es un dígito de control elegido de modo que $\mathbf{c} \cdot \mathbf{u} = 0$ en $\mathbb{Z}_{10}$, donde el vector de control **c** es $[3, 1, 3, 1, 3, 1, 3, 1, 3, 1, 3, 1]$. Esto es, después de reordenar,

$$3(u_1 + u_3 + u_5 + u_7 + u_9 + u_{11}) + (u_2 + u_4 + u_6 + u_8 + u_{10}) + d = 0$$



Figura 1.80
Un código universal de producto

donde d es el dígito de control. En otras palabras, el dígito de control se elige de modo que el lado izquierdo de esta expresión sea un múltiplo de 10.

Para el UPC que se muestra en la figura 1.80, puede determinar que el dígito de control es 6, y realizar todos los cálculos en $\mathbb{Z}_{10}$:

$$\begin{aligned}\mathbf{c} \cdot \mathbf{u} &= 3 \cdot 0 + 7 + 3 \cdot 4 + 9 + 3 \cdot 2 + 7 + 3 \cdot 0 + 2 + 3 \cdot 0 + 9 + 3 \cdot 4 + d \\ &= 3(0 + 4 + 2 + 0 + 0 + 4) + (7 + 9 + 7 + 2 + 9) + d \\ &= 3(0) + 4 + d \\ &= 4 + d\end{aligned}$$

El dígito de control d debe ser 6 para que el resultado del cálculo sea 0 en $\mathbb{Z}_{10}$.

(Otra forma de considerar el dígito de control en este ejemplo es que se elija de modo que $\mathbf{c} \cdot \mathbf{u}$ sea un *múltiplo* de 10.)

El código universal de producto detectará todos los errores individuales y la mayoría de los errores de transposición en componentes adyacentes. Para ver este último punto, suponga que el UPC en el ejemplo 1.40 se escribió de manera incorrecta como $\mathbf{u}' = [0, 7, 4, 2, 9, 7, 0, 2, 0, 9, 4, 6]$, con el cuarto y quinto componentes transpuestos. Cuando se aplica el vector de control, se tendría $\mathbf{c} \cdot \mathbf{u}' = \mathbf{4} \neq 0$ (¡verifíquelo!), lo que alerta del hecho de que hubo un error. (Vea los ejercicios 29 y 30.)

Ejemplo 1.41

El código de número internacional normalizado de libros de 10 dígitos (ISBN-10) es otro código de dígito de control muy utilizado. Está diseñado para detectar más tipos de errores que el código universal de productos y, en consecuencia, es ligeramente más complicado. Aunque el principio básico es el mismo.

El vector código es un vector en $\mathbb{Z}_{11}^{10}$. Los primeros nueve componentes proporcionan el país, editor e información del libro; el décimo componente es el dígito de control. Suponga que el ISBN-10 de un libro es 0-534-34450-X. Se registra como el vector

$$\mathbf{b} = [0, 5, 3, 4, 3, 4, 4, 5, 0, X]$$

donde el "dígito" de control es la letra X.

Para el código ISBN-10, el vector de control es $\mathbf{c} = [10, 9, 8, 7, 6, 5, 4, 3, 2, 1]$, y se requiere que $\mathbf{c} \cdot \mathbf{b} = 0$ en $\mathbb{Z}_{11}$. Determine el dígito de control para el vector **b** en este ejemplo. Debe calcular

$$\mathbf{c} \cdot \mathbf{b} = 10 \cdot 0 + 9 \cdot 5 + 8 \cdot 3 + 7 \cdot 4 + 6 \cdot 3 + 5 \cdot 4 + 4 \cdot 4 + 3 \cdot 5 + 2 \cdot 0 + d$$

donde d es el dígito de control. Comience por realizar todas las multiplicaciones en $\mathbb{Z}_{11}$. (Por ejemplo, $9 \cdot 5 = 1$, pues 45 es 1 más que el menor múltiplo más cercano de 11, a saber, 44. En un reloj de 11 horas, las 45 en punto es 1 en punto.) La suma simplificada es

$$0 + 1 + 2 + 6 + 7 + 9 + 5 + 4 + 0 + d$$

y al sumar en $\mathbb{Z}_{11}$ se obtiene $1 + d$. El dígito de control d debe elegirse ahora de modo que el resultado final sea 0; por tanto, en $\mathbb{Z}_{11}$, $d = 10$. (De manera equivalente, d debe elegirse de modo que $\mathbf{c} \cdot \mathbf{b}$ será un múltiplo de 11.) Pero, dado que es preferible que cada componente de un ISBN-10 sea un solo dígito, el numeral romano X se usa para 10 siempre que funcione como dígito de control, como sucede aquí.

El código ISBN detectará todos los errores individuales y errores de transposición adyacente (vea los ejercicios 33-35).

Comentario El código ISBN-10 se usa desde 1970. Sin embargo, desde 2007, la mayoría de los libros también se identifican con un código ISBN de 13 dígitos. Este dígito ISBN-13 es compatible con el código EAN-13 (número de artículo europeo de 13 dígitos) y usa un esquema de dígito de control similar al UPC. De manera específica, un código ISBN-13 es un vector en $\mathbb{Z}_{10}^{13}$ donde el último dígito es el dígito de control y el vector de control es $[1, 3, 1, 3, \ldots, 3, 1]$ en $\mathbb{Z}_{10}^{13}$. Como su primo UPC, el código ISBN-13 detectará todos los errores individuales, mas no todos los errores de transposición adyacentes.

Ejercicios 1.4

Vectores fuerza

En los ejercicios 1-6, determine la resultante de las fuerzas dadas.

1. $\mathbf{f}_1$ actúa hacia el norte con una magnitud de 12 N y $\mathbf{f}_2$ actúa hacia el este con una magnitud de 5 N
2. $\mathbf{f}_1$ actúa hacia el oeste con una magnitud de 15 N y $\mathbf{f}_2$ actúa hacia el sur con una magnitud de 20 N
3. $\mathbf{f}_1$ actúa con una magnitud de 8 N y $\mathbf{f}_2$ actúa a un ángulo de 60° con respecto a $\mathbf{f}_1$ con una magnitud de 8 N

4. $\mathbf{f}_1$ actúa con una magnitud de 4 N y $\mathbf{f}_2$ actúa a un ángulo de 135° con respecto a $\mathbf{f}_2$ con una magnitud de 6 N
5. $\mathbf{f}_1$ actúa hacia el este con una magnitud de 2 N, $\mathbf{f}_2$ actúa hacia el oeste con una magnitud de 6 N y $\mathbf{f}_3$ actúa a un ángulo de 60° con respecto a $\mathbf{f}_1$ con una magnitud de 4 N
6. $\mathbf{f}_1$ actúa hacia el este con una magnitud de 10 N, $\mathbf{f}_2$ actúa hacia el norte con una magnitud de 13 N, $\mathbf{f}_3$ actúa hacia el oeste con una magnitud de 5 N y $\mathbf{f}_4$ actúa hacia el sur con una magnitud de 8 N
7. Descomponga una fuerza de 10 N en dos fuerzas mutuamente perpendiculares, de modo que un componente forme un ángulo de 60° con la fuerza de 10 N.
8. Un bloque de 10 kg se encuentra sobre una rampa que está inclinada a un ángulo de 30° (figura 1.81). Si supone que no hay fricción, ¿qué fuerza, paralela a la rampa, debe aplicarse para evitar que el bloque se deslice por la rampa?

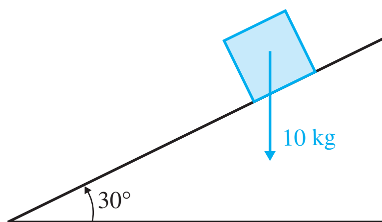


Figura 1.81

9. Una grúa levanta un automóvil. La tensión en el cable de la grúa es 1500 N y el cable forma un ángulo de 45° con la horizontal, como se muestra en la figura 1.82. ¿Cuál es la fuerza vertical que tiende a levantar al automóvil del suelo?

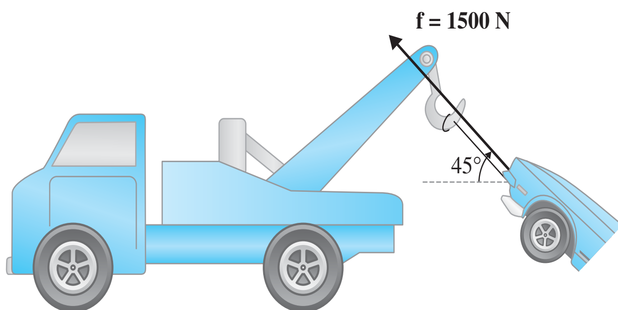


Figura 1.82

10. Una podadora tiene una masa de 30 kg. Se le empuja con una fuerza de 100 N. Si el manubrio de la podadora forma un ángulo de 45° con el suelo, ¿cuál es el componente horizontal de la fuerza que hace que la podadora se mueva hacia adelante?

11. Un letrero que cuelga afuera del restaurante de Joe tiene una masa de 50 kg (figura 1.83). Si el cable que lo soporta forma un ángulo de 60° con el muro del edificio, determine la tensión en el cable.

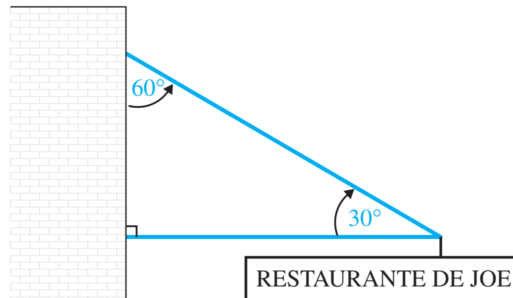


Figura 1.83

12. Un letrero que cuelga en la ventana del restaurante de Joe tiene una masa de 1 kg. Si las cuerdas que lo sostienen forman cada una un ángulo de 45° con el letrero y los ganchos que las sostienen están a la misma altura (figura 1.84), encuentre la tensión en cada cuerda.

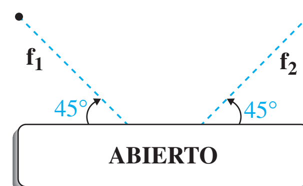


Figura 1.84

13. Una pintura, con una masa de 15 kg, está suspendida mediante dos cables de unos ganchos en el techo. Si los cables tienen longitudes de 15 cm y 20 cm, y la distancia entre los ganchos es 25 cm, encuentre la tensión en cada cable.
14. Una pintura, con una masa de 20 kg, está suspendida mediante dos cables de un techo. Si los cables forman ángulos de 30° y 45° con el techo, encuentre la tensión en cada cable.

Vectores código

En los ejercicios 15 y 16, encuentre el vector código de control de paridad para el vector binario **u**.

15. $\mathbf{u} = [1, 0, 1, 1]$ **16.** $\mathbf{u} = [1, 1, 0, 1, 1]$

En los ejercicios 17-29, se proporciona un vector código de control de paridad **v**. *Determine si en la transmisión de* **v** *ocurrió un error individual.*

17. $\mathbf{v} = [1, 0, 1, 0]$ **18.** $\mathbf{v} = [1, 1, 1, 0, 1, 1]$
19. $\mathbf{v} = [0, 1, 0, 1, 1, 1]$ **20.** $\mathbf{v} = [1, 1, 0, 1, 0, 1, 1, 1]$

Los ejercicios 21-24 se refieren a códigos de dígito de control en los que el vector de control es $\mathbf{c} = [1, 1, \ldots, 1]$ *de la longitud adecuada. En cada caso, encuentre el dígito de control d que se agregaría al vector u.*

21. $\mathbf{u} = [1, 2, 2, 2]$ en $\mathbb{Z}_3^4$ **22.** $\mathbf{u} = [3, 4, 2, 3]$ en $\mathbb{Z}_5^4$

23. $\mathbf{u} = [1, 5, 6, 4, 5]$ en $\mathbb{Z}_7^5$ **24.** $\mathbf{u} = [3, 0, 7, 5, 6, 8]$ en $\mathbb{Z}_9^6$

25. Demuestre que, para cualesquiera enteros positivos m y n, el código de dígito de control en $\mathbb{Z}_m^n$, con vector de control $\mathbf{c} = \mathbf{1} = [1, 1, \ldots, 1]$, detectará todos los errores individuales. (Esto es: demuestre que, si los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{Z}_m^n$ difieren en exactamente una entrada, entonces $\mathbf{c} \cdot \mathbf{u} \neq \mathbf{c} \cdot \mathbf{v}$.)

En los ejercicios 26 y 27, encuentre el dígito de control d en el código universal de producto dado.

26. $[0, 5, 9, 4, 6, 4, 7, 0, 0, 2, 7, d]$

27. $[0, 1, 4, 0, 1, 4, 1, 8, 4, 1, 2, d]$

28. Considere el UPC $[0, 4, 6, 9, 5, 6, 1, 8, 2, 0, 1, 5]$.
 (a) Demuestre que este UPC no puede ser correcto.
 (b) Si supone que se cometió un error individual y que el dígito incorrecto es el 6 en la tercera entrada, encuentre el UPC correcto.

29. Demuestre que el código universal de producto detectará todos los errores individuales.

30. **(a)** Pruebe que si se comete un error de transposición en la segunda y tercera entradas del UPC $[0, 7, 4, 9, 2, 7, 0, 2, 0, 9, 4, 6]$, se detectará el error.
 (b) Demuestre que existe una transposición que involucra dos entradas adyacentes del UPC en el inciso (a) que no se detectaría.
 (c) En general, ¿en qué caso el código universal de productos no detectará un error de transposición que involucre dos entradas adyacentes?

En los ejercicios 31 y 32, encuentre el dígito de control d en el número internacional normalizado de libros (ISBN-10).

30. $[0, 3, 8, 7, 9, 7, 9, 9, 3, d]$

31. $[0, 3, 9, 4, 7, 5, 6, 8, 2, d]$

32. Considere el ISBN-10 $[0, 4, 4, 9, 5, 0, 8, 3, 5, 6]$.
 (a) Demuestre que este ISBN-10 no puede ser correcto.
 (b) Si supone que se cometió un error individual y que el dígito incorrecto es el 5 en la quinta entrada, encuentre el ISBN-10 correcto.

34. **(a)** Pruebe que si se cometió un error de transposición en la cuarta y quinta entradas del ISBN-10 $[0, 6, 7, 9, 7, 6, 2, 9, 0, 6]$ se detectará el error.
 (b) Pruebe que si se cometió un error de transposición en cualesquiera dos entradas adyacentes del ISBN-10 del inciso (a) el error se detectará.
 (c) Pruebe, en general, que el código ISBN-10 siempre detectará un error de transposición que involucre dos entradas adyacentes.

35. Considere el ISBN-10 $[0, 8, 3, 7, 0, 9, 9, 0, 2, 6]$.
 (a) Demuestre que este ISBN-10 no puede ser correcto.
 (b) Si supone que el error fue una transposición que involucra dos entradas adyacentes, encuentre el ISBN-10 correcto.
 (c) Proporcione un ejemplo de un ISBN-10 para el cual se detectará un error de transposición que involucre dos entradas adyacentes, pero no será corregible.

El sistema Codabar

Toda tarjeta de crédito y débito tiene una identificación única mediante un número de 16 dígitos que representa un código de dígito de control. Como en los ejemplos de esta sección, los primeros 15 dígitos los asigna el banco emisor de la tarjeta, mientras que el último es un dígito de control determinado por una fórmula que usa aritmética modular.

Todos los grandes bancos usan un sistema llamado Codabar para asignar el dígito de control. Es una ligera variación sobre el método del código universal de producto y se basa en un algoritmo diseñado por el científico de cómputo de IBM, Hans Peter Luhn en la década de 1950.



Suponga que los primeros 15 dígitos de su tarjeta son

5412 3456 7890 432

y que el dígito de control es d. Esto corresponde al vector

$$\mathbf{x} = [5, 4, 1, 2, 3, 4, 5, 6, 7, 8, 9, 0, 4, 3, 2, d]$$

en $\mathbb{Z}_{10}^{16}$. El sistema Codabar usa el vector de control $\mathbf{c} = [2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1, 2, 1]$, pero en lugar de requerir que $\mathbf{c} \cdot \mathbf{x} = 0$ en $\mathbb{Z}_{10}$, se agrega un cálculo adicional para aumentar la capacidad de detección de error del código. Sea h el conteo del número de dígitos en *posiciones impares* que son *mayores que 4*. En este ejemplo, dichos dígitos son 5, 5, 7 y 9, de modo que $h = 4$.

Ahora se requiere que $\mathbf{c} \cdot \mathbf{x} + h = 0$ en $\mathbb{Z}_{10}$. Por tanto, en el ejemplo se tiene, al reordenar y trabajar módulo 10,

$$\begin{aligned}
\mathbf{c} \cdot \mathbf{x} + h &= (2 \cdot 5 + 4 + 2 \cdot 1 + 2 + 2 \cdot 3 + 4 + 2 \cdot 5 + 6 + 2 \cdot 7 + 8 + 2 \cdot 9 \\
&\qquad + 0 + 2 \cdot 4 + 3 + 2 \cdot 2 + d) + 4 \\
&= 2\,(5 + 1 + 3 + 5 + 7 + 9 + 4 + 2) + (4 + 2 + 4 + 6 + 8 + 0 \\
&\qquad + 3 + d) + 4 \\
&= 2\,(6) + 7 + d + 4 \\
&= 3 + d
\end{aligned}$$

En consecuencia, el dígito de control d para esta tarjeta debe ser 7, de modo que el resultado del cálculo es 0 en $\mathbb{Z}_{10}$.

El sistema Codabar es uno de los métodos más eficientes de detección de error. Detectará todos los errores de dígito individual y la mayoría de otros errores comunes, como los de transposición adyacente.

Repaso del capítulo

Definiciones y conceptos clave

Preguntas de repaso

1. Marque cada uno de los siguientes enunciados como verdadero o falso:

(a) Para los vectores **u**, **v** y **w** en $\mathbb{R}^n$, si $\mathbf{u} + \mathbf{w} = \mathbf{v} + \mathbf{w}$ entonces $\mathbf{u} = \mathbf{v}$.

(b) Para los vectores **u**, **v** y **w** en $\mathbb{R}^n$, si $\mathbf{u} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w}$ entonces $\mathbf{u} = \mathbf{v}$.

(c) Para los vectores **u**, **v** y **w** en $\mathbb{R}^3$, si **u** es ortogonal a **v**, y **v** es ortogonal a **w**, entonces **u** es ortogonal a **w**.

(d) En $\mathbb{R}^3$, si una recta ℓ es paralela a un plano $\mathcal{P}$, entonces un vector director **d** para ℓ es paralelo a un vector normal **n** para $\mathcal{P}$.

(e) en $\mathbb{R}^3$, si una recta ℓ es perpendicular a un plano $\mathcal{P}$, entonces un vector director **d** para ℓ es paralelo a un vector normal **n** para $\mathcal{P}$.

(f) En $\mathbb{R}^3$, si dos planos no son paralelos, entonces deben intersectarse en una recta.

(g) En $\mathbb{R}^3$, si dos rectas no son paralelas, entonces deben intersectarse en un punto.

(h) Si **v** es un vector binario tal que $\mathbf{v} \cdot \mathbf{v} = 0$, entonces $\mathbf{v} = \mathbf{0}$.

(i) En $\mathbb{Z}_5$, si $ab = 0$, entonces $a = 0$ o $b = 0$.

(j) En $\mathbb{Z}_6$, si $ab = 0$, entonces $a = 0$ o $b = 0$.

2. Si $\mathbf{u} = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ y el vector $4\mathbf{u} + \mathbf{v}$ se dibuja con su origen en el punto $(10, -10)$, encuentre las coordenadas del punto en la punta de $4\mathbf{u} + \mathbf{v}$.

3. Si $\mathbf{u} = \begin{bmatrix} -1 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ y $2\mathbf{x} + \mathbf{u} = 3(\mathbf{x} - \mathbf{v})$, calcule **x**.

4. Sean A, B, C y D los vértices de un cuadrado con centro en el origen O, ordenados en el sentido de las manecillas del reloj. Si $\mathbf{a} = \overrightarrow{OA}$ y $\mathbf{b} = \overrightarrow{OB}$, encuentre $\overrightarrow{BC}$ en términos de **a** y **b**.

5. Encuentre el ángulo entre los vectores $[-1, 1, 2]$ y $[2, 1, -1]$.

6. Encuentre la proyección de $\mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ sobre $\mathbf{u} = \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}$.

7. Encuentre un vector unitario en el plano xy que sea ortogonal a $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

8. Encuentre la ecuación general del plano que pasa por el punto $(1, 1, 1)$ y es perpendicular a la recta con ecuaciones paramétricas

$$\begin{aligned} x &= 2 - t \\ y &= 3 + 2t \\ z &= -1 + t \end{aligned}$$

9. Encuentre la ecuación general del plano que pasa por el punto $(3, 2, 5)$ y es paralelo al plano cuya ecuación general es $2x + 3y - z = 0$.

10. Encuentre la ecuación general del plano que pasa por los puntos $A(1, 1, 0)$, $B(1, 0, 1)$ y $C(0, 1, 2)$.

11. Encuentre el área del triángulo con vértices en $A(1, 1, 0)$, $B(1, 0, 1)$ y $C(0, 1, 2)$.

12. Encuentre el punto medio del segmento de recta entre $A = (5, 1, -2)$ y $B = (3, -7, 0)$.

13. ¿Por qué no hay vectores **u** y **v** en $\mathbb{R}^n$ tales que $\|\mathbf{u}\| = 2$, $\|\mathbf{v}\| = 3$ y $\mathbf{u} \cdot \mathbf{v} = -7$?

14. Encuentre la distancia desde el punto $(3, 2, 5)$ hasta el plano cuya ecuación general es $2x + 3y - z = 0$.

15. Encuentre la distancia desde el punto $(3, 2, 5)$ hasta la recta con ecuaciones paramétricas $x = t, y = 1 + t, z = 2 + t$.

16. Calcule $3 - (2 + 4)^3(4 + 3)^2$ en $\mathbb{Z}_5$.

17. Si es posible, resuelva $3(x + 2) = 5$ en $\mathbb{Z}_7$.

18. Si es posible, resuelva $3(x + 2) = 5$ en $\mathbb{Z}_9$.

19. Calcule $[2, 1, 3, 3] \cdot [3, 4, 4, 2]$ en $\mathbb{Z}_5^4$.

20. Sea $\mathbf{u} = [1, 1, 1, 0]$ en $\mathbb{Z}_2^4$. ¿Cuántos vectores binarios **v** satisfacen $\mathbf{u} \cdot \mathbf{v} = 0$?

2

Sistemas de ecuaciones lineales

El mundo está lleno de ecuaciones. . . Debe haber una respuesta para todo, si sólo supiéramos cómo plantear las ecuaciones.

—Anne Tyler
The Accidental Tourist
Alfred A. Knopf, 1985, p. 235

2.0 Introducción: trivialidad

La palabra *trivial* se deriva de la raíz latina *tri* ("tres") y de la palabra latina *via* ("camino"). Por ende, literalmente, una trivialidad es un lugar donde se juntan tres caminos. Este punto de reunión común da lugar al otro significado, más popular, de *trivial*: lugar común, ordinario o insignificante. En las universidades medievales, el *trivium* consistía de las tres materias "comunes" (gramática, retórica y lógica) que se impartían antes del *quadrivium* (aritmética, geometría, música y astronomía). Los "tres caminos" que constituían el trivium fueron el comienzo de las artes liberales.

En esta sección comenzará el examen de los sistemas de ecuaciones lineales. El mismo sistema de ecuaciones puede verse en tres formas diferentes, aunque igualmente importantes; esos serán sus tres caminos, que conducirán todos a la misma solución. Necesitará acostumbrarse a este camino triple de ver los sistemas de ecuaciones lineales, para que se conviertan en lugar común (¡trivial!) para usted.

El sistema de ecuaciones que se considerará es

$$\begin{aligned} 2x + y &= 8 \\ x - 3y &= -3 \end{aligned}$$

Problema 1 Dibuje las dos rectas representadas por dichas ecuaciones. ¿Cuál es su punto de intersección?

Problema 2 Considere los vectores $\mathbf{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$. Dibuje la cuadrícula coordenada determinada por $\mathbf{u}$ y $\mathbf{v}$. [*Sugerencia:* dibuje primero, ligeramente, la cuadrícula coordenada estándar y úsela como auxiliar para dibujar la nueva.]

Problema 3 En la cuadrícula u-v, encuentre las coordenadas de $\mathbf{w} = \begin{bmatrix} 8 \\ -3 \end{bmatrix}$.

Problema 4 Otra forma de enunciar el problema 3 es pedir los coeficientes de x y y para los que $x\mathbf{u} + y\mathbf{v} = \mathbf{w}$. Escriba las dos ecuaciones a las cuales es equivalente esta ecuación vectorial (una para cada componente). ¿Qué observa?

Problema 5 Ahora regrese a las rectas que dibujó para el problema 1. A la recta cuya ecuación es $2x + y = 8$ se le referirá como recta 1, y a la recta cuya ecuación es $x - 3y = -3$ como recta 2. Trace el punto $(0, 0)$ sobre su gráfica para el problema 1 y etiquételo P_0. Di-

Tabla 2.1

Punto	x	y
P_0	0	0
P_1		
P_2		
P_3		
P_4		
P_5		
P_6		

buje un segmento de recta *horizontal* de P_0 a la recta 1 y etiquete este nuevo punto P_2. A continuación, dibuje un segmento de recta *vertical* de P_1 a la recta 2 y etiquete este punto P_2. Ahora dibuje un segmento de recta *horizontal* de P_2 a la recta 1 para obtener el punto P_3. Continúe de esta forma y dibuje segmentos verticales a la recta 2 seguidos por segmentos horizontales a la recta 1. ¿Qué parece estar ocurriendo?

Problema 6 Con una calculadora con dos lugares decimales de precisión, encuentre las coordenadas (aproximadas) de los puntos $P_1, P_2, P_3, \ldots, P_6$. (Descubrirá que es útil despejar primero la primera ecuación para x en términos de y y la segunda ecuación para y en términos de x.) Registre sus resultados en la tabla 2.1 y escriba por separado las coordenadas x y y de cada punto.

Los resultados de estos problemas muestran que la tarea de "resolver" un sistema de ecuaciones lineales puede verse en varias formas. Repita el proceso descrito en los problemas con los siguientes sistemas de ecuaciones:

(a) $\begin{aligned} 4x - 2y &= 0 \\ x + 2y &= 5 \end{aligned}$ (b) $\begin{aligned} 3x + 2y &= 9 \\ x + 3y &= 10 \end{aligned}$ (c) $\begin{aligned} x + y &= 5 \\ x - y &= 3 \end{aligned}$ (d) $\begin{aligned} x + 2y &= 4 \\ 2x - y &= 3 \end{aligned}$

¿Todas sus observaciones de los problemas 1-6 todavía son válidas para estos ejemplos? Note cualquier similitud o diferencia. En este capítulo se explorarán estas ideas con más detalle.

2.1 Introducción a los sistemas de ecuaciones lineales

Recuerde que la ecuación general de una recta en $\mathbb{R}^2$ es de la forma

$$ax + by = c$$

y que la ecuación general de un plano en $\mathbb{R}^3$ es de la forma

$$ax + by + cz = d$$

Las ecuaciones de esta forma se llaman ***ecuaciones lineales***.

Definición Una ***ecuación lineal*** en las n variables $x_1, x_2, \ldots, x_n$ es una ecuación que puede escribirse en la forma

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

donde los ***coeficientes*** $a_1, a_2, \ldots, a_n$ y el ***término constante*** b son constantes.

Ejemplo 2.1

Las siguientes ecuaciones son lineales:

$$3x - 4y = -1 \qquad r - \tfrac{1}{2}s - \tfrac{15}{3}t = 9 \qquad x_1 + 5x_2 = 3 - x_3 + 2x_4$$

$$\sqrt{2}x + \frac{\pi}{4}y - \left(\operatorname{sen}\frac{\pi}{5}\right)z = 1 \qquad 3.2x_1 - 0.01x_2 = 4.6$$

Observe que la tercera ecuación es lineal porque puede reescribirse en la forma $x_1 + 5x_2 + x_3 - 2x_4 = 3$. También es importante notar que, aunque en estos ejemplos (y en la mayoría de las aplicaciones) los coeficientes y términos constantes son números reales, en algunos ejemplos y aplicaciones serán números complejos o miembros de $\mathbb{Z}_p$ para algún número primo p.

Las siguientes ecuaciones no son lineales:

$$xy + 2z = 1 \qquad x_1^2 - x_2^3 = 3 \qquad \frac{x}{y} + z = 2$$

$$\sqrt{2x} + \frac{\pi}{4}y - \operatorname{sen}\left(\frac{\pi}{5}z\right) = 1 \qquad \operatorname{sen}x_1 - 3x_2 + 2^{x_3} = 0$$

Por tanto, las ecuaciones lineales no contienen productos, recíprocos u otras funciones de las variables; las variables sólo aparecen a la primera potencia y sólo se multiplican por constantes. Ponga particular atención al cuarto ejemplo en cada lista: ¿por qué la cuarta ecuación en la primera lista es lineal, pero la cuarta ecuación en la segunda lista no lo es?

Una ***solución*** de una ecuación lineal $a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$ es un vector $[s_1, s_2, \ldots, s_n]$ cuyos componentes satisfacen la ecuación cuando se sustituye $x_1 = s_1$, $x_2 = s_2, \ldots, x_n = s_n$

Ejemplo 2.2

(a) $[5, 4]$ es una solución de $x - 4y = -1$ porque, cuando se sustituye $x = 5$ y $y = 4$, la ecuación se satisface: $3(5) - 4(4) = -1$. $[1, 1]$ es otra solución. En general, las soluciones simplemente corresponden a los puntos sobre la recta determinada por la ecuación dada. Por tanto, al hacer $x = t$ y despejar y, se ve que el conjunto completo de soluciones puede escribirse en la forma paramétrica $[t, \frac{1}{4} + \frac{3}{4}t]$. (También podría hacer y igual a algún parámetro, por decir, s, y despejar x; las dos soluciones paramétricas se verían diferentes, pero serían equivalentes. Inténtelo.)

(b) La ecuación lineal $x_1 - x_2 + 2x_3 = 3$ tiene $[3, 0, 0]$, $[0, 1, 2]$ y $[6, 1, -1]$ como soluciones específicas. El conjunto completo de soluciones corresponde al conjunto de puntos en el plano determinado por la ecuación dada. Si se hace $x_2 = s$ y $x_3 = t$, entonces una solución paramétrica está dada por $[3 + s - 2t, s, t]$.(¿Cuáles valores de s y t producen las tres soluciones específicas anteriores?)

Un ***sistema de ecuaciones lineales*** es un conjunto finito de ecuaciones lineales, cada una con las mismas variables. Una ***solución*** de un sistema de ecuaciones lineales es un vector que *simultáneamente* es una solución de cada ecuación en el sistema. El ***conjunto solución*** de un sistema de ecuaciones lineales es el conjunto de *todas* las soluciones del sistema. Al proceso de encontrar el conjunto solución de un sistema de ecuaciones lineales se le conocerá como "resolver el sistema".

Ejemplo 2.3

El sistema

$$\begin{aligned} 2x - y &= 3 \\ x + 3y &= 5 \end{aligned}$$

tiene $[2,1]$ como una solución, pues es una solución de ambas ecuaciones. Por otra parte, $[1, -1]$ no es una solución del sistema, porque sólo satisface la primera ecuación.

Ejemplo 2.4

Resuelva los siguientes sistemas de ecuaciones lineales:

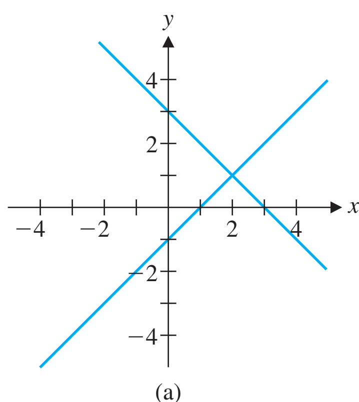
(a) $\begin{aligned} x - y &= 1 \\ x + y &= 3 \end{aligned}$ (b) $\begin{aligned} x - y &= 2 \\ 2x - 2y &= 4 \end{aligned}$ (c) $\begin{aligned} x - y &= 1 \\ x - y &= 3 \end{aligned}$

Solución

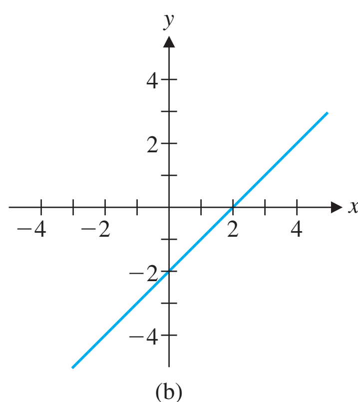
(a) Al sumar las dos ecuaciones se obtiene $2x = 4$, de modo que $x = 2$, de donde se encuentra que $y = 1$. Una comprobación rápida confirma que $[2, 1]$ es de hecho una solución de ambas ecuaciones. Que ésta es la *única* solución que puede verse al observar que esta solución corresponde al (único) punto de intersección $(2, 1)$ de las rectas con ecuaciones $x - y = 1$ y $x + y = 3$, como se muestra en la figura 2.1(a). Por tanto, $[2, 1]$ es una *solución única*.

(b) La segunda ecuación en este sistema es justo el doble de la primera, de modo que las soluciones corresponden sólo a la primera ecuación: a saber, los puntos sobre la recta $x - y = 2$. Las mismas pueden representarse paramétricamente como $[2 + t, t]$. Por tanto, este sistema tiene *un número infinito de soluciones* [figura 2.1(b)].

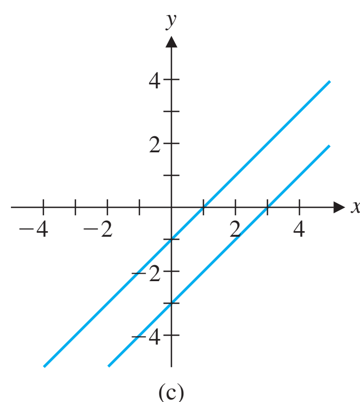
(c) Dos números x y y no pueden tener simultáneamente una diferencia de 1 y 3. Por tanto, este sistema *no tiene soluciones*. (Un planteamiento más algebraico puede ser restar la segunda ecuación de la primera, lo que produce la conclusión absurda $0 = -2$.) Como muestra la figura 2.1(c), en este caso las rectas para las ecuaciones son paralelas.



(a)



(b)



(c)

Figura 2.1

Un sistema de ecuaciones lineales se llama ***consistente*** si tiene al menos una solución. Un sistema sin soluciones se llama ***inconsistente***. Aun cuando sean pequeños, los tres sistemas en el ejemplo 2.4 ilustran las únicas tres posibilidades para el número de soluciones de un sistema de ecuaciones lineales con coeficientes reales. Más tarde se probará que estas mismas tres posibilidades se sostienen para *cualquier* sistema de ecuaciones lineales sobre los números reales.

Un sistema de ecuaciones lineales con coeficientes reales tiene

(a) una solución única (un sistema consistente) o
(b) un número infinito de soluciones (un sistema consistente) o
(c) ninguna solución (un sistema inconsistente).

Resolución de un sistema de ecuaciones lineales

Dos sistemas lineales se llaman ***equivalentes*** si tienen los mismos conjuntos solución. Por ejemplo,

$$\begin{aligned} x - y &= 1 \\ x + y &= 3 \end{aligned} \qquad \text{y} \qquad \begin{aligned} x - y &= 1 \\ y &= 1 \end{aligned}$$



son equivalentes, pues ambos tienen la solución única $[2, 1]$. (Compruébelo.)

El planteamiento para resolver un sistema de ecuaciones lineales es transformar el sistema dado en uno equivalente que sea más fácil de resolver. El patrón triangular del segundo ejemplo anterior (en el que la segunda ecuación tiene una variable menos que la primera) es el que se buscará.

Ejemplo 2.5 Resuelva el sistema

$$\begin{aligned} x - y - z &= 2 \\ y + 3z &= 5 \\ 5z &= 10 \end{aligned}$$

Solución A partir de la última ecuación y trabajando hacia atrás, se descubre sucesivamente que $z = 2$, $y = 5 - 3(2) = -1$ y $x = 2 + (-1) + 2 = 3$. De modo que la solución única es $[3, -1, 2]$.

El procedimiento que se usó para resolver el ejemplo 2.5 se llama ***sustitución hacia atrás***.

Ahora se regresa a la estrategia general para transformar un sistema dado en uno equivalente que puede resolverse fácilmente mediante sustitución hacia atrás. Este proceso se describirá con mayor detalle en la siguiente sección; por ahora, simplemente se le observará en acción en un solo ejemplo.

Ejemplo 2.6 Resuelva el sistema

$$\begin{aligned} x - y - z &= 2 \\ 3x - 3y + 2z &= 16 \\ 2x - y + z &= 9 \end{aligned}$$

Solución Para transformar este sistema en uno que muestre la estructura triangular del ejemplo 2.5, primero es necesario eliminar la variable x de las ecuaciones 2 y 3. Observe que restar de las ecuaciones 2 y 3 múltiplos adecuados de la ecuación 1 hará el truco. A continuación, observe que se opera sobre los coeficientes, no sobre las variables, de modo que puede ahorrar cierta escritura si registra los coeficientes y términos constantes en la *matriz*

$$\left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right]$$

La palabra *matriz* se deriva de la palabra latina *mater*, que significa "madre". Cuando se agrega el sufijo *–ix*, el significado se convierte en "útero". Así como un útero rodea un feto, los corchetes de una matriz rodean sus entradas, y así como el útero da origen a un bebé, una matriz da origen a ciertos tipos de funciones llamadas *transformaciones lineales*. Una matriz con m renglones y n columnas se llama matriz de $m \times n$ (pronúnciese "m por n").

donde las primeras tres columnas contienen los coeficientes de las variables en orden, la columna final contiene los términos constantes y la barra vertical sirve para recordar los signos igual en las ecuaciones. Esta matriz se llama ***matriz aumentada*** del sistema.

Existen varias formas de convertir el sistema dado en uno con el patrón triangular que se busca. Los pasos que se usarán aquí son los más cercanos en espíritu al método más general que se describe en la siguiente sección. Se realizará la secuencia de operaciones sobre un sistema dado y simultáneamente sobre la matriz aumentada correspondiente. Comience por eliminar x de las ecuaciones 2 y 3.

$$\begin{aligned} x - y - z &= 2 \\ 3x - 3y + 2z &= 16 \\ 2x - y + z &= 9 \end{aligned} \qquad \left[\begin{array}{rrr|r} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right]$$

Reste 3 veces la primera ecuación de la segunda ecuación:

$$\begin{aligned} x - y - z &= 2 \\ 5z &= 10 \\ 2x - y + z &= 9 \end{aligned}$$

Reste 3 veces el primer renglón del segundo renglón:

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 2 & -1 & 1 & 9 \end{array}\right]$$

Reste 2 veces la primera ecuación de la tercera ecuación:

$$\begin{aligned} x - y - z &= 2 \\ 5z &= 10 \\ y + 3z &= 5 \end{aligned}$$

Reste 2 veces el primer renglón del tercer renglón:

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 0 & 1 & 3 & 5 \end{array}\right]$$

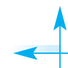
Intercambie las ecuaciones 2 y 3:

$$\begin{aligned} x - y - z &= 2 \\ y + 3z &= 5 \\ 5z &= 10 \end{aligned}$$

Intercambie los renglones 2 y 3:

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{array}\right]$$

Éste es el mismo sistema que se resolvió usando la sustitución hacia atrás en el ejemplo 2.5, donde se descubrió que la solución era [3, –1, 2]. Por tanto, ésta también es la solución al sistema dado en este ejemplo. ¿Por qué? El cálculo anterior muestra que *cualquier solución del sistema dado también es una solución del sistema final*, pero, dado que los pasos recién realizados son *reversibles*, podría recuperar el sistema original a partir del sistema final. (¿Cómo?) Por tanto, *cualquier solución del sistema final también es una solución del sistema dado.* Por ende, los sistemas son equivalentes (como lo son todos los obtenidos en los pasos intermedios anteriores).  Más aún, se puede trabajar con matrices en lugar de ecuaciones, pues es asunto simple reinsertar las variables antes de proceder con la sustitución hacia atrás. (Trabajar con matrices es el tema de la siguiente sección.)



Comentario Las calculadoras con capacidades matriciales y los sistemas algebraicos de cómputo (CAS por sus siglas en inglés) pueden facilitar la resolución de sistemas de ecuaciones lineales, en particular cuando los sistemas son grandes o tienen coeficientes que no son "bonitos", como suele ser el caso en aplicaciones de la vida real. Sin embargo, como siempre, debe hacer tantos ejemplos como pueda con lápiz y papel hasta que se sienta cómodo con las técnicas. Incluso si se solicita una calculadora o CAS, piense en *cómo* realizaría los cálculos de forma manual antes de hacer algo. Después de obtener una respuesta, asegúrese de pensar si es razonable.

No se confunda al pensar que la tecnología siempre le dará la respuesta más rápido o más fácilmente que los cálculos a mano. ¡En ocasiones pueden no darle la respuesta en absoluto! Los errores de redondeo asociados con la aritmética de punto flotante que usan las calculadoras y computadoras pueden causar serios problemas y conducir a respuestas totalmente erróneas a ciertos problemas. Vea la Exploración "Mentiras que me dice mi computadora" para un vistazo del problema. (¡Está advertido!)

Ejercicios 2.1

En los ejercicios 1-6, determine cuáles son ecuaciones lineales con las variables x, y y z. Si alguna ecuación no es lineal, explique por qué.

1. $x - \pi y + \sqrt[3]{5}z = 0$

2. $x^2 + y^2 + z^2 = 1$

3. $x^{-1} + 7y + z = \operatorname{sen}\left(\frac{\pi}{9}\right)$

4. $2x - xy - 5z = 0$

5. $3\cos x - 4y + z = \sqrt{3}$

6. $(\cos 3)x - 4y + z = \sqrt{3}$

En los ejercicios 7-10, encuentre una ecuación lineal que tenga el mismo conjunto solución que la ecuación dada (posiblemente con algunas restricciones sobre las variables).

7. $2x + y = 7 - 3y$

8. $\dfrac{x^2 - y^2}{x - y} = 1$

9. $\dfrac{1}{x} + \dfrac{1}{y} = \dfrac{4}{xy}$

10. $\log_{10} x - \log_{10} y = 2$

En los ejercicios 11-14, encuentre el conjunto solución de cada ecuación.

11. $3x - 6y = 0$

12. $2x_1 + 3x_2 = 5$

13. $x + 2y + 3z = 4$

14. $4x_1 + 3x_2 + 2x_3 = 1$

En los ejercicios 15-18, dibuje gráficas que correspondan a los sistemas lineales dados. Determine geométricamente si cada sistema tiene una solución única, un número infinito de soluciones o ninguna solución. Luego resuelva cada sistema algebraicamente para confirmar su respuesta.

15. $\begin{aligned} x + y &= 0 \\ 2x + y &= 3 \end{aligned}$

16. $\begin{aligned} x - 2y &= 7 \\ 3x + y &= 7 \end{aligned}$

17. $\begin{aligned} 3x - 6y &= 3 \\ -x + 2y &= 1 \end{aligned}$

18. $\begin{aligned} 0.10x - 0.05y &= 0.20 \\ -0.06x + 0.03y &= -0.12 \end{aligned}$

En los ejercicios 19-24, resuelva el sistema dado mediante sustitución hacia atrás.

19. $\begin{aligned} x - 2y &= 1 \\ y &= 3 \end{aligned}$

20. $\begin{aligned} 2u - 3v &= 5 \\ 2v &= 6 \end{aligned}$

21. $\begin{aligned} x - y + z &= 0 \\ 2y - z &= 1 \\ 3z &= -1 \end{aligned}$

22. $\begin{aligned} x_1 + 2x_2 + 3x_3 &= 0 \\ -5x_2 + 2x_3 &= 0 \\ 4x_3 &= 0 \end{aligned}$

23. $\begin{aligned} x_1 + x_2 - x_3 - x_4 &= 1 \\ x_2 + x_3 + x_4 &= 0 \\ x_3 - x_4 &= 0 \\ x_4 &= 1 \end{aligned}$

24. $\begin{aligned} x - 3y + z &= 5 \\ y - 2z &= -1 \end{aligned}$

Los sistemas en los ejercicios 25 y 26 muestran un patrón "triangular inferior" que las hace fáciles de resolver mediante sustitución hacia atrás. (La sustitución hacia adelante se encontrará nuevamente en el capítulo 3.) Resuelva estos sistemas.

25. $\begin{aligned} x &= 2 \\ 2x + y &= -3 \\ -3x - 4y + z &= -10 \end{aligned}$

26. $\begin{aligned} x_1 &= -1 \\ -\tfrac{1}{2}x_1 + x_2 &= 5 \\ \tfrac{3}{2}x_1 + 2x_2 + x_3 &= 7 \end{aligned}$

Encuentre las matrices aumentadas de los sistemas lineales en los ejercicios 27-30.

27. $\begin{aligned} x - y &= 0 \\ 2x + y &= 3 \end{aligned}$

28. $\begin{aligned} 2x_1 + 3x_2 - x_3 &= 1 \\ x_1 + x_3 &= 0 \\ -x_1 + 2x_2 - 2x_3 &= 0 \end{aligned}$

29. $\begin{aligned} x + 5y &= -1 \\ -x + y &= -5 \\ 2x + 4y &= 4 \end{aligned}$

30. $\begin{aligned} a - 2b + d &= 2 \\ -a + b - c - 3d &= 1 \end{aligned}$

En los ejercicios 31 y 32, encuentre un sistema de ecuaciones lineales que tenga la matriz dada como su matriz aumentada.

31. $\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & -1 & 0 & 1 \\ 2 & -1 & 1 & 1 \end{array}\right]$

32. $\left[\begin{array}{ccccc|c} 1 & -1 & 0 & 3 & 1 & 2 \\ 1 & 1 & 2 & 1 & -1 & 4 \\ 0 & 1 & 0 & 2 & 3 & 0 \end{array}\right]$

Para los ejercicios 33-38, resuelva los sistemas lineales en los ejercicios dados.

33. Ejercicio 27

34. Ejercicio 28

35. Ejercicio 29

36. Ejercicio 30

37. Ejercicio 31

38. Ejercicio 32

39. **(a)** Encuentre un sistema de dos ecuaciones lineales con las variables x y y cuyo conjunto solución esté dado por las ecuaciones paramétricas $x = t$ y $y = 3 - 2t$.

(b) Encuentre otra solución paramétrica para el sistema en el inciso (a) en el que el parámetro es s y $y = s$.

40. **(a)** Encuentre un sistema de dos ecuaciones lineales con las variables x_1, x_2 y x_3 cuyo conjunto solución esté dado por las ecuaciones paramétricas $x_1 = t$, $x_2 = 1 + t$ y $x_3 = 2 - t$.

(b) Encuentre otra solución paramétrica para el sistema en el inciso (a) en el que el parámetro es s y $x_3 = s$.

En los ejercicios 41-44, los sistemas de ecuaciones son no lineales. Encuentre sustituciones (cambios de variables) que conviertan cada sistema en un sistema lineal y use este último para ayudar a resolver el sistema dado.

41. $\dfrac{2}{x} + \dfrac{3}{y} = 0$

$\dfrac{3}{x} + \dfrac{4}{y} = 1$

42. $x^2 + 2y^2 = 6$

$x^2 - y^2 = 3$

43. $\tan x - 2 \operatorname{sen} y = 2$

$\tan x - \operatorname{sen} y + \cos z = 2$

$\operatorname{sen} y - \cos z = -1$

44. $-2^a + 2(3^b) = 1$

$3(2^a) - 4(3^b) = 1$

2.2 Métodos directos para resolver sistemas lineales

En esta sección se verá un procedimiento sistemático general para resolver un sistema de ecuaciones lineales. Este procedimiento se basa en la idea de reducir la matriz aumentada del sistema dado a una forma que luego pueda resolverse mediante sustitución hacia atrás. El método es *directo* en el sentido de que conduce directamente a la solución (si existe una) en un número finito de pasos. En la sección 2.5 se considerarán algunos métodos *indirectos* que funcionan en una forma completamente diferente.

Matrices y forma escalonada

Existen dos importantes matrices asociadas con un sistema lineal. La ***matriz de coeficientes*** contiene los coeficientes de las variables, y la ***matriz aumentada*** (que ya se encontró) es la matriz coeficiente aumentada por una columna adicional que contiene los términos constantes.

Para el sistema

$$\begin{aligned} 2x + y - z &= 3 \\ x \qquad + 5z &= 1 \\ -x + 3y - 2z &= 0 \end{aligned}$$

la matriz de coeficientes es

$$\begin{bmatrix} 2 & 1 & -1 \\ 1 & 0 & 5 \\ -1 & 3 & -2 \end{bmatrix}$$

y la matriz aumentada es

$$\left[\begin{array}{rrr|r} 2 & 1 & -1 & 3 \\ 1 & 0 & 5 & 1 \\ -1 & 3 & -2 & 0 \end{array}\right]$$

Observe que si falta una variable (como y en la segunda ecuación), su coeficiente 0 se ingresa en la posición adecuada de la matriz. Si la matriz de coeficientes de un sistema lineal se denota con A y el vector columna de términos constantes con $\mathbf{b}$, entonces la forma de la matriz aumentada es $[A \mid \mathbf{b}]$.

Al resolver un sistema lineal, no siempre será posible reducir la matriz de coeficientes a la forma triangular, como se hizo en el ejemplo 2.6. Sin embargo, siempre se puede lograr un patrón en escalera en las entradas distintas de cero de la matriz final.

La palabra *escalonada* proviene de la palabra latina *scala*, que significa "escalera". La palabra francesa para "escalera", *échelle*, también se deriva de esta base latina. Una matriz en forma escalonada muestra un patrón en escalera.

Definición Una matriz está en ***forma escalonada por renglones*** si satisface las siguientes propiedades:

1. Cualquier renglón que consiste completamente de ceros está en la parte baja.
2. En cada renglón distinto de cero, el primer elemento distinto de cero (llamado ***elemento pivote***) está en una columna a la izquierda de cualquier elemento pivote bajo él.

Observe que estas propiedades garantizan que los elementos pivote formen un patrón en escalera. En particular, en cualquier columna que contenga un elemento pivote, todas las entradas bajo el elemento pivote son cero, como ilustran los siguientes ejemplos.

Ejemplo 2.7

Las siguientes matrices están en forma escalonada por renglones:

$$\begin{bmatrix} 2 & 4 & 1 \\ 0 & -1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 5 \\ 0 & 0 & 4 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & 2 & 1 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 2 & 0 & 1 & -1 & 3 \\ 0 & 0 & -1 & 1 & 2 & 2 \\ 0 & 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

Si una matriz en forma escalonada por renglones en realidad es la matriz aumentada de un sistema lineal, el sistema es bastante fácil de resolver mediante sustitución hacia atrás solamente.

Ejemplo 2.8

Si supone que cada una de las matrices en el ejemplo 2.7 es una matriz aumentada, escriba los sistemas correspondientes de ecuaciones lineales y resuélvalos.

Solución Recuerde primero que la última columna en una matriz aumentada es el vector de términos constantes. Entonces la primera matriz corresponde al sistema

$$\begin{aligned} 2x_1 + 4x_2 &= 1 \\ -x_2 &= 2 \end{aligned}$$

(Advierta que se eliminó la última ecuación $0 = 0$ o $0x_1 + 0x_2 = 0$, que claramente es satisfecha por cualquier valor de x_1 y x_2.) La sustitución hacia atrás produce $x_2 = -2$ y entonces $2x_1 = 1 - 4(-2) = 9$, de modo que $x_1 = \frac{9}{2}$. La solución es $[\frac{9}{2}, -2]$.

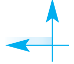
La segunda matriz tiene el sistema correspondiente

$$\begin{aligned} x_1 \quad &= 1 \\ x_2 &= 5 \\ 0 &= 4 \end{aligned}$$

La última ecuación representa $0x_1 + 0x_2 = 4$, que claramente no tiene solución. Por tanto, el sistema no tiene soluciones. De igual modo, el sistema que corresponde a la cuarta matriz no tiene soluciones. Para el sistema que corresponde a la tercera matriz, se tiene

$$\begin{aligned} x_1 + x_2 + 2x_3 &= 1 \\ x_3 &= 3 \end{aligned}$$

de modo que $x_1 = 1 - 2(3) - x_2 = -5 - x_2$. Existe un número infinito de soluciones, pues a x_2 se le puede asignar cualquier valor t para obtener la solución paramétrica $[-5 - t, t, 3]$.



Operaciones elementales con renglones

Ahora se describe el procedimiento mediante el cual cualquier matriz puede reducirse a una matriz en forma escalonada por renglones. Las operaciones permisibles, llamadas ***operaciones elementales con renglones***, corresponden a las operaciones que pueden realizarse sobre un sistema de ecuaciones lineales para transformarlo en un sistema equivalente.

Definición Las siguientes ***operaciones elementales con renglones*** pueden realizarse sobre una matriz:

1. Intercambiar dos renglones.
2. Multiplicar un renglón por una constante distinta de cero.
3. Sumar un múltiplo de un renglón a otro renglón.

Observe que la división de un renglón entre una constante distinta de cero está implícita en la definición anterior pues, por ejemplo, dividir un renglón entre 2 es lo mismo que multiplicarlo por $\frac{1}{2}$. De igual modo, restar un múltiplo en un renglón de otro renglón es lo mismo que sumar un múltiplo negativo de un renglón a otro renglón.

Se usará la siguiente notación abreviada para las tres operaciones elementales con renglones:

1. $R_i \leftrightarrow R_j$ significa intercambiar renglones i y j.
2. kR_i significa multiplicar el renglón i por k.
3. $R_i + kR_j$ significa sumar k veces el renglón j al renglón i (y sustituir el renglón i con el resultado).

El proceso de aplicar operaciones elementales con renglones para llevar una matriz a la forma escalonada por renglones, llamado ***reducción de renglón***, se usa para reducir una matriz a la forma escalonada.

Ejemplo 2.9

Reduzca la siguiente matriz a forma escalonada:

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix}$$

Solución Trabaje columna por columna, de izquierda a derecha y de arriba abajo. La estrategia es crear un elemento pivote en una columna y luego usarlo para crear ceros abajo de él. El elemento elegido para convertirse en elemento pivote se llama simplemente ***pivote***, y esta fase del proceso se conoce como ***pivoteo***. Aunque no es estrictamente necesario, con frecuencia es conveniente usar la segunda operación elemental de renglón para convertir en 1 cada elemento pivote.

Comience por introducir ceros en la primera columna abajo del 1 pivote en el primer renglón:

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix} \xrightarrow[R_4 + R_1]{\substack{R_2 - 2R_1 \\ R_3 - 2R_1}} \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 3 & -1 & 2 & 10 \end{bmatrix}$$

Ahora la primera columna está como se quiere, de modo que la siguiente cosa por hacer es crear un elemento pivote en el segundo renglón, en busca del patrón en escalera de la forma escalonada. En este caso, se hace al intercambiar renglones. (También podría sumar los renglones 3 o 4 al renglón 2.)

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & 3 & -1 & 2 & 10 \end{bmatrix}$$

El pivote esta vez fue -1. Ahora se crea un cero en la parte baja de la columna 2, usando el elemento pivote -1 en el renglón 2:

$$\xrightarrow{R_4 + 3R_2} \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 8 & 8 & -8 \\ 0 & 0 & 29 & 29 & -5 \end{bmatrix}$$

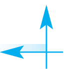
Ahora se hace la columna 2. Al notar que ya se tiene un elemento pivote en la columna 3, sólo se pivotea en el 8 para introducir un cero debajo de él. Esto es más sencillo si primero se divide el renglón 3 entre 8:

$$\xrightarrow{\frac{1}{8}R_3} \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 29 & 29 & -5 \end{bmatrix}$$

Ahora use el 1 pivote en el renglón 3 para crear un cero debajo de él:

$$\xrightarrow{R_4 - 29R_3} \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix}$$

Con este paso final, la matriz se redujo a forma escalonada.



Comentarios

- La forma escalonada por renglones de una matriz no es única. (Encuentre una forma escalonada por renglones diferente para la matriz del ejemplo 2.9.)

- El elemento pivote en cada renglón se usa para crear los ceros debajo de él.
- Los pivotes no necesariamente son los elementos que están originalmente en las posiciones que al final ocupan los elementos pivote. En el ejemplo 2.9, los pivotes fueron 1, −1, 8 y 24. La matriz original tenía 1, 4, 2 y 5 en dichas posiciones sobre la "escalera".
- Una vez que se pivotea y se introducen ceros bajo el elemento pivote en una columna, dicha columna no cambia. En otras palabras, la forma escalonada por renglones surge de izquierda a derecha, de arriba abajo.

Las operaciones elementales con renglones son reversibles; esto es, pueden "deshacerse". Por ende, si alguna operación elemental con renglones convierte A en B, existe también una forma escalonada por renglones que convierte B en A. (Vea los ejercicios 15 y 16.)

Definición Las matrices A y B son ***equivalentes por renglones*** si existe una secuencia de operaciones elementales con renglones que convierta A en B.

Las matrices en el ejemplo 2.9,

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix}$$

son equivalentes por renglones. Sin embargo, en general, ¿cómo puede decirse si dos matrices son equivalentes por renglones?

Teorema 2.1 Las matrices A y B son equivalentes por renglones si y sólo si pueden reducirse a la misma forma escalonada por renglones.

Demostración Si A y B son equivalentes por renglones, entonces más operaciones con renglones reducirán B (y por tanto a A) a la (misma) forma escalonada por renglones.

Recíprocamente, si A y B tienen la misma forma escalonada por renglones R, entonces, mediante operaciones elementales con renglones, puede convertirse A en R y B en R. Al invertir la última secuencia de operaciones, puede convertir R en B, y por tanto la secuencia $A \to R \to B$ logra el efecto deseado.

Comentario En la práctica, el Teorema 2.1 es más fácil de usar si R es la forma escalonada *reducida* por renglones de A y B, como se define en la página 79. Vea los ejercicios 17 y 18.

Eliminación gaussiana

Cuando se aplica la reducción por renglones a la matriz aumentada de un sistema de ecuaciones lineales, se crea un sistema equivalente que puede resolverse mediante sustitución hacia atrás. Todo el proceso se conoce como ***eliminación gaussiana***.

Eliminación gaussiana

1. Escriba la matriz aumentada del sistema de ecuaciones lineales.
2. Use operaciones elementales con renglones para reducir la matriz aumentada a forma escalonada por renglones.
3. Con sustitución hacia atrás, resuelva el sistema equivalente que corresponda a la matriz reducida por renglones.

Comentario Cuando se realiza a mano, el paso 2 de la eliminación gaussiana permite muchas opciones. He aquí algunos lineamientos útiles:

(a) Localice la columna de la extrema izquierda que no sea toda ceros.

(b) Cree un elemento pivote en la parte superior de esta columna. (Por lo general será más sencillo si la convierte en 1 pivote. Vea el ejercicio 22.)

(c) Use el elemento pivote para crear ceros debajo de él.

(d) Cubra el renglón que contiene el elemento pivote y regrese al paso (a) para repetir el procedimiento sobre la submatriz restante. Deténgase cuando toda la matriz esté en forma escalonada por renglones.

Ejemplo 2.10 Resuelva el sistema

$$\begin{aligned} 2x_2 + 3x_3 &= 8 \\ 2x_1 + 3x_2 + x_3 &= 5 \\ x_1 - x_2 - 2x_3 &= -5 \end{aligned}$$

Solución La matriz aumentada es

$$\left[\begin{array}{rrr|r} 0 & 2 & 3 & 8 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array}\right]$$

Proceda a reducir esta matriz a forma escalonada por renglones y siga los lineamientos dados para el paso 2 del proceso. La primera columna distinta de cero es la columna 1.

Carl Friedrich Gauss (1777–1855) es considerado generalmente uno de los tres más grandes matemáticos de todos los tiempos, junto con Arquímedes y Newton. Con frecuencia se le llama "príncipe de los matemáticos", un sobrenombre que merecía con creces. Niño prodigio, Gauss supuestamente era capaz de hacer aritmética antes de poder hablar. A la edad de 3 años, corrigió un error en los cálculos de su padre para la nómina de la compañía, y como joven estudiante descubrió la fórmula $n(n + 1)/2$ para la suma de los primeros n números naturales. Cuando tenía 19, demostró que podía construirse un polígono de 17 lados usando solamente regla y compás, y a los 21 demostró, en su disertación doctoral, que todo polinomio de grado n con coeficientes reales o complejos tenía exactamente n ceros, contando múltiples ceros: el teorema fundamental del álgebra.

La publicación en 1801 de Gauss, *Disquisitiones Arithmeticae*, se considera generalmente el fundamento de la moderna teoría de números, pero hizo aportaciones a casi cada rama de las matemáticas, así como a la estadística, la física, la astronomía y la topografía. Gauss no publicó todos sus hallazgos, probablemente porque era muy crítico de su propio trabajo. Tampoco le gustaba dar clases y con frecuencia criticaba a otros matemáticos, acaso porque descubrió, mas no publicó, sus resultados antes que ellos.

El método llamado eliminación gaussiana lo conocían los chinos en el tercer siglo antes de Cristo, pero lleva el nombre de Gauss debido a su redescubrimiento en un ensayo en el que resolvió un sistema de ecuaciones lineales para describir la órbita de un asteroide.

Comience por crear un elemento pivote en la parte superior de esta columna; e intercambiar los renglones 1 y 3 es la mejor forma de lograr esto.

$$\left[\begin{array}{rrr|r} 0 & 2 & 3 & 8 \\ 2 & 3 & 1 & 5 \\ 1 & -1 & -2 & -5 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{rrr|r} 1 & -1 & -2 & -5 \\ 2 & 3 & 1 & 5 \\ 0 & 2 & 3 & 8 \end{array}\right]$$

Ahora cree un segundo cero en la primera columna, usando el 1 pivote:

$$\xrightarrow{R_2 - 2R_1} \left[\begin{array}{rrr|r} 1 & -1 & -2 & -5 \\ 0 & 5 & 5 & 15 \\ 0 & 2 & 3 & 8 \end{array}\right]$$

Cubra ahora el primer renglón y repita el procedimiento. La segunda columna es la primera columna distinta de cero de la submatriz. Al multiplicar el renglón 2 por $\frac{1}{5}$ creará un 1 pivote.

$$\left[\begin{array}{rrr|r} 1 & -1 & -2 & -5 \\ 0 & 5 & 5 & 15 \\ 0 & 2 & 3 & 8 \end{array}\right] \xrightarrow{\frac{1}{5}R_2} \left[\begin{array}{rrr|r} 1 & -1 & -2 & -5 \\ 0 & 1 & 1 & 3 \\ 0 & 2 & 3 & 8 \end{array}\right]$$

Ahora necesita otro cero en la parte baja de la columna 2:

$$\xrightarrow{R_3 - 2R_2} \left[\begin{array}{rrr|r} 1 & -1 & -2 & -5 \\ 0 & 1 & 1 & 3 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

La matriz aumentada ahora está en forma escalonada por renglones y se avanza al paso 3. El sistema correspondiente es

$$\begin{aligned} x_1 - x_2 - 2x_3 &= -5 \\ x_2 + x_3 &= 3 \\ x_3 &= 2 \end{aligned}$$

y la sustitución hacia atrás produce $x_3 = 3$, entonces $x_2 = 3 - x_3 = 3 - 2 = 1$, y finalmente $x_1 = -5 + x_2 + 2x_3 = -5 + 1 + 4 = 0$. La solución se escribe en forma vectorial como

$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$$

(A partir de ahora, las soluciones vectoriales de los sistemas lineales se escribirán como vectores columna. La razón de esto será clara en el capítulo 3.)

Ejemplo 2.11

Resuelva el sistema

$$\begin{aligned} w - x - y + 2z &= 1 \\ 2w - 2x - y + 3z &= 3 \\ -w + x - y &= -3 \end{aligned}$$

Solución La matriz aumentada es

$$\left[\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{array}\right]$$

que puede reducirse por renglones del modo siguiente:

$$\left[\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 2 & -2 & -1 & 3 & 3 \\ -1 & 1 & -1 & 0 & -3 \end{array}\right] \xrightarrow[R_3+R_1]{R_2-2R_1} \left[\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & 2 & -2 \end{array}\right]$$

$$\xrightarrow{R_3+2R_2} \left[\begin{array}{rrrr|r} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

El sistema asociado ahora es

$$\begin{aligned} w - x - y + 2z &= 1 \\ y - z &= 1 \end{aligned}$$

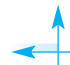
que tiene un número infinito de soluciones. Existe más de una forma de asignar parámetros, pero se procederá a usar sustitución hacia atrás, y escribir las variables correspondientes a los elementos pivote (las ***variables pivote***) en términos de las otras variables (las ***variables libres***).

En este caso, las variables pivote son w y y, y las variables libres son x y z. Por tanto, $y = 1 + z$, y a partir de esto se obtiene

$$\begin{aligned} w &= 1 + x + y - 2z \\ &= 1 + x + (1 + z) - 2z \\ &= 2 + x - z \end{aligned}$$

Si se asignan los parámetros $x = s$ y $z = t$, la solución puede escribirse en forma vectorial como

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 + s - t \\ s \\ 1 + t \\ t \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + s\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$



El ejemplo 2.11 destaca una propiedad muy importante: en un sistema consistente, las variables libres son las que no son variables pivote. Dado que el número de variables pivote es el número de renglones distintos de cero en la forma escalonada por renglones de la matriz de coeficientes, puede predecirse el número de variables libres (parámetros) antes de encontrar la solución explícita con el uso de sustitución hacia atrás. En el capítulo 3 se probará que, aunque la forma escalonada por renglones de una matriz no es única, el número de renglones distintos de cero es el mismo en *todas* las formas escalonadas por renglones de una matriz dada. Por tanto, tiene sentido dar un nombre a este número.

Definición El ***rank*** de una matriz es el número de renglones distintos de cero en su forma escalonada por renglones.

El rank de una matriz A se denotará mediante rank(A). En el ejemplo 2.10, el rank de la matriz de coeficientes es 3, y en el ejemplo 2.11, el rank de la matriz de coeficientes es 2. Las observaciones recién realizadas justifican el siguiente teorema, que se demostrará con más generalidad en los capítulos 3 y 6.

Teorema 2.2 **El teorema del rank**

Sea A la matriz de coeficientes de un sistema de ecuaciones lineales con n variables. Si el sistema es consistente, entonces

$$\text{número de variables libres} = n - \text{rank}(A)$$

Por tanto, en el ejemplo 2.10, se tiene $3 - 3 = 0$ variables libres (en otras palabras, una solución *única*), y en el ejemplo 2.11, se tiene $4 - 2 = 2$ variables libres, como se encontró.

Ejemplo 2.12 Resuelva el sistema

$$\begin{aligned} x_1 - x_2 + 2x_3 &= 3 \\ x_1 + 2x_2 - x_3 &= -3 \\ 2x_2 - 2x_3 &= 1 \end{aligned}$$

Solución Cuando la matriz aumentada se reduce por renglones, se tiene

$$\left[\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 0 & 3 & -3 & -6 \\ 0 & 2 & -2 & 1 \end{array}\right]$$

$$\xrightarrow{\frac{1}{3}R_2} \left[\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 2 & -2 & 1 \end{array}\right]$$

$$\xrightarrow{R_3 - 2R_2} \left[\begin{array}{rrr|r} 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 5 \end{array}\right]$$

lo que conduce a la ecuación imposible $0 = 5$. (También podría hacerse $R_3 - \frac{2}{3}R_2$ como la segunda operación elemental con renglones, lo que daría la misma contradicción pero una forma escalonada por renglones diferente.) Por tanto, el sistema no tiene soluciones: es inconsistente.

Wilhelm Jordan (1842–1899) fue un profesor alemán de geodesia cuyas aportaciones a la resolución de sistemas lineales fue un método sistemático de sustitución hacia atrás estrechamente relacionado con el descrito aquí.

Eliminación de Gauss-Jordan

Una modificación de la eliminación gaussiana simplifica enormemente la fase de sustitución hacia atrás y es particularmente útil cuando se hacen cálculos a mano en un sis-

tema con un número infinito de soluciones. Esta variante, conocida como ***eliminación de Gauss-Jordan***, se apoya en reducir aún más la matriz aumentada.

Definición Una matriz está en ***forma escalonada reducida por renglones*** si satisface las siguientes propiedades:

1. Está en forma escalonada por renglones.
2. El elemento pivote en cada renglón distinto de cero es 1 (llamado ***1 pivote***).
3. Cada columna que contiene un 1 pivote tiene ceros en todos los otros lugares.

La siguiente matriz está en forma escalonada reducida por renglones:

$$\begin{bmatrix} 1 & 2 & 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Para matrices de 2×2, las posibles formas escalonadas reducidas por renglones son

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & * \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ y } \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Para una demostración breve de que la forma escalonada reducida por renglones de una matriz es única, vea el artículo de Thomas Yuster, "The Reduced Row Echelon Form of a Matrix Is Unique: A Simple Proof", en el número de marzo de 1984 de *Mathematics Magazine* (vol. 57, núm. 2, pp. 93-94).

donde ***** puede ser cualquier número.

Es claro que, después de que una matriz se reduce a forma escalonada, posteriores operaciones elementales con renglones la llevarán a la forma escalonada reducida por renglones. Lo que no es claro (aunque la intuición puede sugerirlo) es que, a diferencia de la forma escalonada por renglones, la forma escalonada reducida por renglones de una matriz es *única*.

En la eliminación de Gauss-Jordan, se procede como en la eliminación gaussiana, pero la matriz aumentada se reduce a la forma escalonada *reducida* por renglones.

Eliminación de Gauss-Jordan

1. Escriba la matriz aumentada del sistema de ecuaciones lineales.
2. Use operaciones elementales con renglones para reducir la matriz aumentada a la forma escalonada reducida por renglones.
3. Si el sistema resultante es consistente, resuelva para las variables pivote en términos de cualquier variable libre restante.

Ejemplo 2.13

Resuelva el sistema en el ejemplo 2.11 mediante eliminación de Gauss-Jordan.

Solución La reducción procede como se hizo en el ejemplo 2.11, hasta que se llega a la forma escalonada:

$$\left[\begin{array}{cccc|c} 1 & -1 & -1 & 2 & 1 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Ahora debe crear un cero arriba del 1 pivote en el segundo renglón, tercera columna. Esto se hace al sumar el renglón 2 al renglón 1 para obtener

$$\left[\begin{array}{rrrr|r} 1 & -1 & 0 & 1 & 2 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Ahora el sistema se reduce a

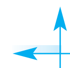
$$\begin{aligned} w - x \qquad + z &= 2 \\ y - z &= 1 \end{aligned}$$

Ahora es mucho más fácil resolver para las variables pivote:

$$w = 2 + x - z \qquad \text{y} \qquad y = 1 + z$$

Si se asignan los parámetros $x = s$ y $z = t$ como antes, la solución puede escribirse en forma vectorial como

$$\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2 + s - t \\ s \\ 1 + t \\ t \end{bmatrix}$$



Comentario Desde un punto de vista computacional, es más eficiente (en el sentido de que requiere menos cálculos) primero reducir la matriz a forma escalonada por renglones y luego, *de derecha a izquierda*, convertir cada elemento pivote a 1 y crear ceros arriba de dichos 1 pivote. Sin embargo, para cálculo manual, descubrirá que es más sencillo sólo trabajar de izquierda a derecha y crear los 1 pivote y los ceros en sus columnas conforme avance.

Regrese a la geometría que lo llevó a este punto. Así como los sistemas de ecuaciones lineales en dos variables corresponden a rectas en $\mathbb{R}^2$, del mismo modo las ecuaciones lineales en tres variables corresponden a planos en $\mathbb{R}^3$. De hecho, muchas preguntas acerca de rectas y planos pueden responderse al resolver un sistema lineal adecuado.

Ejemplo 2.14

Encuentre la recta de intersección de los planos $x + 2y - z = 3$ y $2x + 3y + z = 1$.

Solución Primero, observe que *habrá* una recta de intersección, pues los vectores normales de los dos planos, $[1, 2, -1]$ y $[2, 3, 1]$, no son paralelos. Los puntos que yacen en la intersección de los dos planos corresponden a los puntos en el conjunto solución del sistema

$$\begin{aligned} x + 2y - z &= 3 \\ 2x + 3y + z &= 1 \end{aligned}$$

Aplicar la eliminación de Gauss-Jordan a la matriz aumentada produce

$$\left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 2 & 3 & 1 & 1 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 2 & -1 & 3 \\ 0 & -1 & 3 & -5 \end{array}\right]$$

$$\xrightarrow[-R_2]{R_1 + 2R_2} \left[\begin{array}{ccc|c} 1 & 0 & 5 & -7 \\ 0 & 1 & -3 & 5 \end{array}\right]$$

Al sustituir variables se tiene

$$\begin{aligned} x \quad + 5z &= -7 \\ y - 3z &= 5 \end{aligned}$$

La variable libre z se iguala al parámetro t y por tanto se obtienen las ecuaciones paramétricas de la recta de intersección de los dos planos:

$$\begin{aligned} x &= -7 - 5t \\ y &= 5 + 3t \\ z &= t \end{aligned}$$

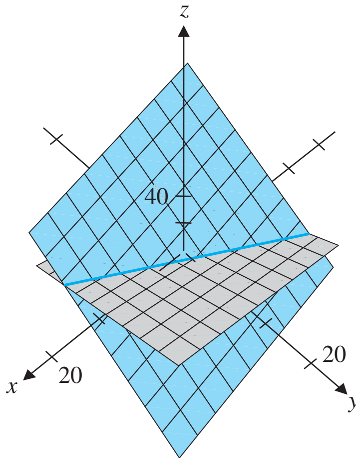


Figura 2.2
La intersección de dos planos

En forma vectorial, la ecuación es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -7 \\ 5 \\ 0 \end{bmatrix} + t\begin{bmatrix} -5 \\ 3 \\ 1 \end{bmatrix}$$

Vea la figura 2.2.

Ejemplo 2.15

Sea $\mathbf{p} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\mathbf{q} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}$. Determine si las rectas $\mathbf{x} = \mathbf{p} + t\mathbf{u}$ y $\mathbf{x} = \mathbf{q} + t\mathbf{v}$ se intersectan y, si es así, encuentre su punto de intersección.

Solución Debe tener cuidado. Aunque t se usó como el parámetro en las ecuaciones de ambas rectas, éstas son independientes y, por tanto, lo son sus parámetros. Use un parámetro diferente (por decir, s) para la primera recta, de modo que su ecuación se convierte en $\mathbf{x} = \mathbf{p} + s\mathbf{u}$. Si las rectas se intersectan, entonces se quiere encontrar $\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que satisfaga ambas ecuaciones simultáneamente. Esto es, se quiere $\mathbf{x} = \mathbf{p} + s\mathbf{u} = \mathbf{q} + t\mathbf{v}$ o $s\mathbf{u} - t\mathbf{v} = \mathbf{q} - \mathbf{p}$.

Al sustituir los $\mathbf{p}$, $\mathbf{q}$, $\mathbf{u}$ y $\mathbf{v}$ dados, se obtienen las ecuaciones

$$\begin{aligned} s - 3t &= -1 \\ s + t &= 2 \\ s + t &= 2 \end{aligned}$$

cuya solución se encuentra fácilmente que es $s = \frac{5}{4}$, $t = \frac{3}{4}$. Por tanto, el punto de intersección es

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \frac{5}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{9}{4} \\ \frac{5}{4} \\ \frac{1}{4} \end{bmatrix}$$

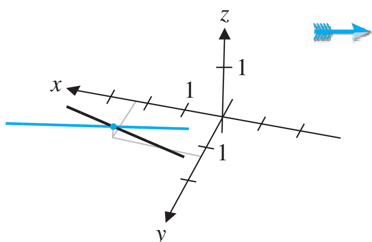


Figura 2.3
Dos rectas que se intersectan

Vea la figura 2.3. (Compruebe que al sustituir $t = \frac{3}{4}$ en la otra ecuación produce el mismo punto.)

Comentario En $\mathbb{R}^3$, es posible que dos rectas se intersecten en un punto, que sean paralelas o ninguna de las dos. Las rectas no paralelas que no se intersectan se llaman *rectas alabeadas*.

Sistemas homogéneos

Se vio que todo sistema de ecuaciones lineales no tiene soluciones tiene una solución única o tiene un número infinito de soluciones. Sin embargo, hay un tipo de sistema que siempre tiene al menos una solución.

Definición Un sistema de ecuaciones lineales se denomina ***homogéneo*** si el término constante en cada ecuación es cero.

En otras palabras, un sistema homogéneo tiene una matriz aumentada de la forma $[A \mid \mathbf{0}]$. El siguiente sistema es homogéneo:

$$\begin{aligned} 2x + 3y - z &= 0 \\ -x + 5y + 2z &= 0 \end{aligned}$$

Dado que un sistema homogéneo no puede no tener solución (¡disculpe la doble negación!), tendrá o una solución única (a saber, la solución cero o trivial) o un número infinito de soluciones. El siguiente teorema dice que el último caso *debe* ocurrir si el número de variables es mayor que el número de ecuaciones.

Teorema 2.3 Si $[A \mid \mathbf{0}]$ es un sistema homogéneo de m ecuaciones lineales con n variables, donde $m < n$, entonces el sistema tiene un número infinito de soluciones.

Demostración Dado que el sistema tiene al menos la solución cero, es consistente. Además, $\operatorname{rank}(A) \leq m$ (¿por qué?). Por el teorema del rank se tiene

$$\text{número de variables libres} = n - \operatorname{rank}(A) \geq n - m > 0$$

De modo que existen al menos una variable libre y, por tanto, un número infinito de soluciones.

Nota El Teorema 2.3 no dice algo acerca del caso donde $m \geq n$. El ejercicio 44 le pide dar ejemplos para demostrar que, en este caso, puede haber una solución única o un número infinito de soluciones.

Sistemas lineales sobre $\mathbb{Z}_p$

$\mathbb{R}$ y $\mathbb{Z}_p$ son ejemplos de *campos*. El conjunto de los números racionales $\mathbb{Q}$ y el conjunto de los números complejos $\mathbb{C}$ son otros ejemplos. Los campos se estudian con detalle en los cursos de álgebra abstracta.

Hasta el momento, todos los sistemas lineales encontrados involucraron números reales, y en concordancia las soluciones han sido vectores en algún $\mathbb{R}^n$. Se vio cómo surgen otros sistemas numéricos, notablemente $\mathbb{Z}_p$. Cuando p es un número primo, $\mathbb{Z}_p$ se comporta en muchas aspectos como $\mathbb{R}$; en particular, es posible sumar, restar, multiplicar y dividir (por números distintos de cero). Por ende, también pueden resolverse sistemas de ecuaciones lineales cuando las variables y coeficientes pertenecen a algún $\mathbb{Z}_p$. En tal caso, se hace referencia a resolver un sistema *sobre* $\mathbb{Z}_p$.

Por ejemplo, la ecuación lineal $x_1 + x_2 + x_3 = 1$, cuando se ve como una ecuación sobre $\mathbb{Z}_2$, tiene exactamente cuatro soluciones:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

(donde la última solución surge porque $1 + 1 + 1 = 1$ en $\mathbb{Z}_2$).

Cuando la ecuación $x_1 + x_2 + x_3 = 1$ se ve sobre $\mathbb{Z}_3$, las soluciones $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ son

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$


(Compruebe esto.)

Pero no es necesario usar métodos de ensayo y error; la reducción por renglones de las matrices aumentadas funciona igual de bien sobre $\mathbb{Z}_p$ que sobre $\mathbb{R}$.

Ejemplo 2.16 Resuelva el siguiente sistema de ecuaciones lineales sobre $\mathbb{Z}_3$.

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 0 \\ x_1 \qquad + x_3 &= 2 \\ x_2 + 2x_3 &= 1 \end{aligned}$$

Solución La primera cosa a notar en ejemplos como este es que no se necesitan resta y división; puede lograr los mismos efectos al usar suma y multiplicación. (Sin embargo, esto requiere que comience a trabajar sobre $\mathbb{Z}_p$, donde p es un primo; vea el ejercicio 60 al final de esta sección y el ejercicio 57 en la sección 1.1.)

Reduzca por renglones la matriz aumentada del sistema y use cálculos en módulo 3.

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right] \xrightarrow{R_2 + 2R_1} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right]$$

$$\xrightarrow[R_3 + 2R_2]{R_1 + R_2} \left[\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 2 & 2 \end{array}\right]$$

$$\xrightarrow[2R_3]{R_1+R_3} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Por tanto, la solución es $x_1 = 1, x_2 = 2, x_3 = 1$.

Ejemplo 2.17 Resuelva el siguiente sistema de ecuaciones lineales sobre $\mathbb{Z}_2$:

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 &= 1 \\ x_1 + x_2 &= 1 \\ x_2 + x_3 &= 0 \\ x_3 + x_4 &= 0 \\ x_1 + x_4 &= 1 \end{aligned}$$

Solución La reducción por renglones procede del modo siguiente:

$$\left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \end{array}\right] \xrightarrow[R_5+R_1]{R_2+R_1} \left[\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[\substack{R_1 + R_2 \\ R_5 + R_2}]{R_2 \leftrightarrow R_3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow[R_4+R_3]{R_2+R_3} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Por tanto, se tiene

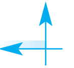
$$\begin{aligned} x_1 + x_4 &= 1 \\ x_2 + x_4 &= 0 \\ x_3 + x_4 &= 0 \end{aligned}$$

Al hacer la variable libre $x_4 = t$ produce

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 + t \\ t \\ t \\ t \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Dado que t puede tomar los dos valores 0 y 1, existen exactamente dos soluciones:

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$





Comentario Para sistemas lineales sobre $\mathbb{Z}_p$, nunca puede haber un número infinito de soluciones. (¿Por qué no?) En vez de ello, cuando hay más de una solución, el número de soluciones es finito y es función del número de variables libres y p. (Vea el ejercicio 59.)

Ejercicios 2.2

En los ejercicios 1-8, determine si la matriz dada está en forma escalonada por renglones. Si lo está, indique si también está en forma escalonada reducida por renglones.

1. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix}$

2. $\begin{bmatrix} 7 & 0 & 1 & 0 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

3. $\begin{bmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

4. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 0 & 3 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 5 & 0 & 1 \end{bmatrix}$

6. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

8. $\begin{bmatrix} 2 & 1 & 3 & 5 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 9-14, use operaciones elementales con renglones para reducir la matriz dada a (a) forma escalonada por renglones y (b) forma escalonada reducida por renglones.

9. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

10. $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$

11. $\begin{bmatrix} 3 & 5 \\ 5 & -2 \\ 2 & 4 \end{bmatrix}$

12. $\begin{bmatrix} 2 & -4 & -2 & 6 \\ 3 & -6 & 2 & 6 \end{bmatrix}$

13. $\begin{bmatrix} 3 & -2 & -1 \\ 2 & -1 & -1 \\ 4 & -3 & -1 \end{bmatrix}$

14. $\begin{bmatrix} -2 & 6 & -7 \\ 3 & -9 & 10 \\ 1 & -3 & 3 \end{bmatrix}$

15. Invierta las operaciones elementales con renglones usadas en el ejemplo 2.9 para demostrar que puede convertir

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 0 & -1 & 10 & 9 & -5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 24 \end{bmatrix} \text{ en}$$

$$\begin{bmatrix} 1 & 2 & -4 & -4 & 5 \\ 2 & 4 & 0 & 0 & 2 \\ 2 & 3 & 2 & 1 & 5 \\ -1 & 1 & 3 & 6 & 5 \end{bmatrix}$$

16. En general, ¿cuál es la operación elemental con renglones que "deshace" cada una de las tres operaciones elementales con renglones $R_i \leftrightarrow R_j$, kR_i y $R_i + kR_j$?

En los ejercicios 17 y 18, demuestre que las matrices dadas son equivalentes por renglones y encuentre una secuencia de operaciones elementales con renglones que convertirán A en B.

17. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 3 & -1 \\ 1 & 0 \end{bmatrix}$

18. $A = \begin{bmatrix} 2 & 0 & -1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 3 & 1 & -1 \\ 3 & 5 & 1 \\ 2 & 2 & 0 \end{bmatrix}$

19. ¿Qué está mal con la siguiente "demostración" de que toda matriz con al menos dos renglones es equivalente por renglones a una matriz con un renglón cero?

Realice $R_2 + R_1$ y $R_1 + R_2$. Ahora los renglones 1 y 2 son idénticos. Ahora realice $R_2 - R_1$ para obtener un renglón de ceros en el segundo renglón.

20. ¿Cuál es el efecto neto de realizar la siguiente secuencia de operaciones elementales con renglones sobre una matriz (con al menos dos renglones)?

$$R_2 + R_1, R_1 - R_2, R_2 + R_1, -R_1$$

21. Los estudiantes frecuentemente realizan el siguiente tipo de cálculo para introducir un cero en una matriz:

$$\begin{bmatrix} 3 & 1 \\ 2 & 4 \end{bmatrix} \xrightarrow{3R_2 - 2R_1} \begin{bmatrix} 3 & 1 \\ 0 & 10 \end{bmatrix}$$

Sin embargo, $3R_2 - 2R_1$ *no* es una operación elemental con renglones. ¿Por qué no? demuestre cómo lograr el mismo resultado usando operaciones elementales con renglones.

22. Considere la matriz $A = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$. Demuestre que cualquiera de los tres tipos de operaciones elementales con renglones puede usarse para crear un 1 pivote en la parte superior de la primera columna. ¿Cuál prefiere y por qué?

23. ¿Cuál es el rank de cada una de las matrices en los ejercicios 1-8?

24. ¿Cuáles son las posibles formas escalonadas reducidas por renglones de las matrices de 3×3?

En los ejercicios 25-34, resuelva el sistema de ecuaciones dado usando eliminación gaussiana o de Gauss-Jordan.

25. $\begin{aligned} x_1 + 2x_2 - 3x_3 &= 9 \\ 2x_1 - x_2 + x_3 &= 0 \\ 4x_1 - x_2 + x_3 &= 4 \end{aligned}$

26. $\begin{aligned} x + 2y &= -1 \\ 2x + y + z &= 1 \\ -x + y - z &= -1 \end{aligned}$

27. $\begin{aligned} x_1 - 3x_2 - 2x_3 &= 0 \\ -x_1 + 2x_2 + x_3 &= 0 \\ 2x_1 + 4x_2 + 6x_3 &= 0 \end{aligned}$

28. $\begin{aligned} 3w + 3x + y &= 1 \\ 2w + x + y + z &= 1 \\ 2w + 3x + y - z &= 2 \end{aligned}$

29. $\begin{aligned} 2r + s &= 3 \\ 4r + s &= 7 \\ 2r + 5s &= -1 \end{aligned}$

30. $\begin{aligned} -x_1 + 3x_2 - 2x_3 + 4x_4 &= 2 \\ 2x_1 - 6x_2 + x_3 - 2x_4 &= -1 \\ x_1 - 3x_2 + 4x_3 - 8x_4 &= -4 \end{aligned}$

31. $\begin{aligned} \tfrac{1}{2}x_1 + x_2 - x_3 - 6x_4 &= 2 \\ \tfrac{1}{6}x_1 + \tfrac{1}{2}x_2 - 3x_4 + x_5 &= -1 \\ \tfrac{1}{3}x_1 - 2x_3 - 4x_5 &= 8 \end{aligned}$

32. $\begin{aligned} \sqrt{2}x + y + 2z &= 1 \\ \sqrt{2}y - 3z &= -\sqrt{2} \\ -y + \sqrt{2}z &= 1 \end{aligned}$

33. $\begin{aligned} w + x + 2y + z &= 1 \\ w - x - y + z &= 0 \\ x + y &= -1 \\ w + x + z &= 2 \end{aligned}$

34. $\begin{aligned} a + b + c + d &= 10 \\ a + 2b + 3c + 4d &= 30 \\ a + 3b + 6c + 10d &= 65 \\ a + 4b + 8c + 15d &= 93 \end{aligned}$

En los ejercicios 35-38, determine por inspección (esto es: sin realizar cálculo alguno) si un sistema lineal con la matriz aumentada dada tiene una solución única, un número infinito de soluciones o ninguna solución. Justifique sus respuestas.

35. $\left[\begin{array}{ccc|c} 0 & 0 & 1 & 2 \\ 0 & 1 & 3 & 1 \\ 1 & 0 & 1 & 1 \end{array}\right]$

36. $\left[\begin{array}{cccc|c} 3 & -2 & 0 & 1 & 1 \\ 1 & 2 & -3 & 1 & -1 \\ 2 & 4 & -6 & 2 & 0 \end{array}\right]$

37. $\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 0 \\ 5 & 6 & 7 & 8 & 0 \\ 9 & 10 & 11 & 12 & 0 \end{array}\right]$

38. $\left[\begin{array}{ccccc|c} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 2 & 1 \\ 7 & 7 & 7 & 7 & 7 & 7 \end{array}\right]$

12. Demuestre que, si $ad - bc \neq 0$, entonces el sistema

$$\begin{aligned} ax + by &= r \\ cx + dy &= s \end{aligned}$$

tiene una solución única.

En los ejercicios 40-43, ¿para qué valor(es) de k, si hay alguno, los sistemas tendrán (a) ninguna solución, (b) una solución única y (c) un número infinito de soluciones?

40. $kx + y = -2$
$2x - 2y = 4$

41. $x + ky = 1$
$kx + y = 1$

42. $x + y + z = 2$
$x + 4y - z = k$
$2x - y + 4z = k^2$

43. $x + y + kz = 1$
$x + ky + z = 1$
$kx + y + z = -2$

44. Proporcione ejemplos de sistemas homogéneos de m ecuaciones lineales con n variables con $m = n$ y con $m > n$ que tengan (a) un número infinito de soluciones y (b) una solución única.

En los ejercicios 45 y 46, encuentre la recta de intersección de los planos dados.

45. $3x + 2y + z = -1$ y $2x - y + 4z = 5$

46. $4x + y - z = 0$ y $2x - y + 3z = 4$

47. **(a)** Proporcione un ejemplo de tres planos que tengan una recta de intersección común (figura 2.4).

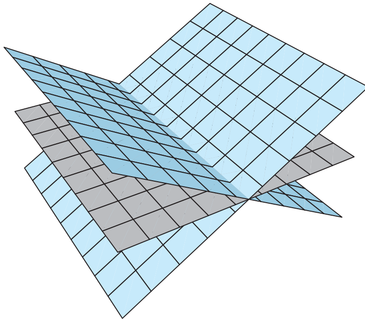


Figura 2.4

(b) Proporcione ejemplos de tres planos que se intersecten en pares, pero que no tengan un punto común de intersección (figura 2.5).

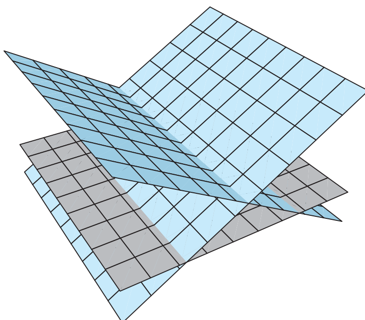


Figura 2.5

(c) Proporcione un ejemplo de tres planos, exactamente dos de los cuales son paralelos (figura 2.6).

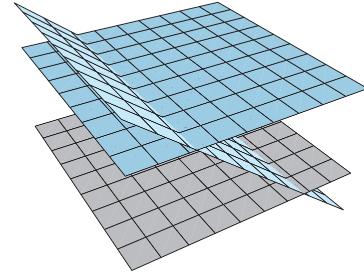


Figura 2.6

(d) Proporcione un ejemplo de tres planos que se intersecten en un solo punto (figura 2.7).

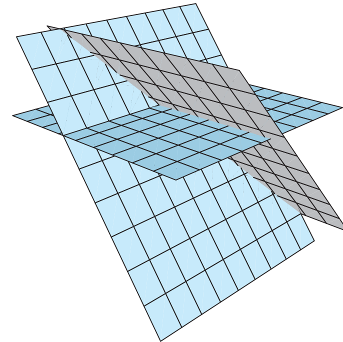


Figura 2.7

En los ejercicios 48 y 49, determine si las rectas $\mathbf{x} = \mathbf{p} + s\mathbf{u}$ *y* $\mathbf{x} = \mathbf{q} + t\mathbf{v}$ *se intersectan y, si lo hacen, encuentre su punto de intersección.*

48. $\mathbf{p} = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$

49. $\mathbf{p} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \mathbf{q} = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$

50. Sean $\mathbf{p} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$. Describa todos los puntos $Q = (a, b, c)$ tales que la recta a través de Q con vector director $\mathbf{v}$ intersecta la recta con la ecuación $\mathbf{x} = \mathbf{p} + s\mathbf{u}$.

51. Recuerde que el producto cruz de dos vectores $\mathbf{u}$ y $\mathbf{v}$ es un vector $\mathbf{u} \times \mathbf{v}$ que es ortogonal tanto a $\mathbf{u}$ como a $\mathbf{v}$. (Vea la Exploración: "el producto cruz" en el capítulo 1.) Si

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

demuestre que existe un número infinito de vectores

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

que simultáneamente satisfacen $\mathbf{u} \cdot \mathbf{x} = 0$ y $\mathbf{v} \cdot \mathbf{x} = 0$ y que todos son múltiplos de

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$$

52. Sean $\mathbf{p} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{q} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 0 \\ 6 \\ -1 \end{bmatrix}$.

Demuestre que las rectas $\mathbf{x} = \mathbf{p} + s\mathbf{u}$ y $\mathbf{x} = \mathbf{q} + t\mathbf{v}$ son rectas oblícuas. Encuentre las ecuaciones vectoriales de un par de planos paralelos, uno que contenga a cada recta.

En los ejercicios 53-58, resuelva los sistemas de ecuaciones lineales sobre el $\mathbb{Z}_p$ indicado.

53. $\begin{aligned} x + 2y &= 1 \\ x + y &= 2 \end{aligned}$ sobre $\mathbb{Z}_3$

54. $\begin{aligned} x + y \phantom{{}+z} &= 1 \\ y + z &= 0 \\ x \phantom{{}+y} + z &= 1 \end{aligned}$ sobre $\mathbb{Z}_2$

55. $\begin{aligned} x + y \phantom{{}+z} &= 1 \\ y + z &= 0 \\ x \phantom{{}+y} + z &= 1 \end{aligned}$ sobre $\mathbb{Z}_3$

56. $\begin{aligned} 3x + 2y &= 1 \\ x + 4y &= 1 \end{aligned}$ sobre $\mathbb{Z}_5$

57. $\begin{aligned} 3x + 2y &= 1 \\ x + 4y &= 1 \end{aligned}$ sobre $\mathbb{Z}_7$

58. $\begin{aligned} x_1 \phantom{{}+2x_2+4x_3} + 4x_4 &= 1 \\ x_1 + 2x_2 + 4x_3 \phantom{{}+4x_4} &= 3 \\ 2x_1 + 2x_2 \phantom{{}+4x_3} + x_4 &= 1 \\ x_1 \phantom{{}+2x_2} + 3x_3 \phantom{{}+4x_4} &= 2 \end{aligned}$ sobre $\mathbb{Z}_5$

59. Demuestre el siguiente corolario del teorema del rank: sea A una matriz de $m \times n$ con elementos en $\mathbb{Z}_p$. Cualquier sistema consistente de ecuaciones lineales con matriz de coeficientes A tiene exactamente $p^{n-\text{rank}(A)}$ soluciones sobre $\mathbb{Z}_p$.

60. Cuando p no es primo, se necesita cuidado adicional para resolver un sistema lineal (o, de hecho, cualquier ecuación) sobre $\mathbb{Z}_p$. Con la eliminación gaussiana, resuelva el siguiente sistema sobre $\mathbb{Z}_6$ ¿Qué complicaciones surgen?

$$\begin{aligned} 2x + 3y &= 4 \\ 4x + 3y &= 2 \end{aligned}$$

Exploración

CAS

Mentiras que me dice mi computadora

Las computadoras y calculadoras almacenan números reales en *formato de punto flotante*. Por ejemplo, 2001 se almacena como 0.2001×10^4, y -0.00063 se almacena como -0.63×10^{-3}. En general, el formato de punto flotante de un número es $\pm M \times 10^k$, donde k es un entero y la *mantisa* M es un número real (decimal) que satisface $0.1 \leq M < 1$.

El número máximo de lugares decimales que pueden almacenarse en la mantisa dependen de la computadora, calculadora o sistema de álgebra computacional. Si el número máximo de lugares decimales que pueden almacenarse es d, se dice que existen d *dígitos significativos*. Muchas calculadoras almacenan 8 o 12 dígitos significativos; las computadoras pueden almacenar más, pero todavía está sujeto a un límite. Cualquier dígito que no se almacene se omite (en cuyo caso se dice que los números están *truncados*) o se usa para *redondear* el número a d dígitos significativos.

Por ejemplo, $\pi \approx 3.141592654$, y su formato en punto flotante es 0.3141592654×10^1. En una computadora que trunca a cinco dígitos significativos, π se almacenaría como 0.31415×10^1 (y se mostraría como 3.1415); una computadora que redondee a cinco dígitos significativos almacenaría π como 0.31416×10^1 (y mostraría 3.1416). Cuando el dígito eliminado es un solitario 5, el último dígito restante se redondea, de modo que se vuelve par. Por ende, redondeado a dos dígitos significativos, 0.735 se convierte en 0.74, mientras que 0.725 se convierte en 0.72.

Siempre que ocurre truncamiento o redondeo, se introduce un *error de redondeo*, que puede tener un efecto dramático sobre los cálculos. Mientras más operaciones se realicen, más errores se acumulan. En ocasiones, por desgracia, no hay nada que se pueda hacer. Esta exploración ilustra este fenómeno con sistemas muy simples de ecuaciones lineales.

1. Resuelva el siguiente sistema de ecuaciones lineales *exactamente* (esto es: trabaje con números racionales a lo largo de los cálculos).

$$\begin{aligned} x + y &= 0 \\ x + \tfrac{801}{800}y &= 1 \end{aligned}$$

2. Como decimal, $\frac{801}{800} = 1.00125$, de modo que, redondeado a cinco dígitos significativos, el sistema se convierte en

$$\begin{aligned} x + y &= 0 \\ x + 1.0012y &= 1 \end{aligned}$$

Con una calculadora o CAS, resuelva este sistema y redondee el resultado de cada cálculo a cinco dígitos significativos.

3. Resuelva el sistema dos veces más, y primero redondee a cuatro dígitos significativos y luego a tres dígitos significativos. ¿Qué ocurre?

4. Claramente, un error de redondeo muy pequeño (menor o igual a 0.00125) puede resultar en errores muy grandes en la solución. Explique geométricamente por qué. (Considere las gráficas de los diversos sistemas lineales que resolvió en los problemas 1-3.)

Los sistemas como el recién trabajado se llaman ***mal condicionados***. Son extremadamente sensibles a errores de redondeo y no hay mucho que se pueda hacer al respecto. En los capítulos 3 y 7 se encontrarán nuevamente sistemas mal condicionados. He aquí otro ejemplo para experimentar:

$$4.552x + 7.083y = 1.931$$
$$1.731x + 2.693y = 2.001$$

Juguetee con varias cantidades de dígitos significativos para ver qué ocurre, a partir de ocho dígitos significativos (si puede).

Pivoteo parcial

En la Exploración "Mentiras que me dice mi computadora", se vio que los sistemas lineales mal condicionados causan problemas cuando ocurren errores de redondeo. En esta exploración descubrirá otra forma en la que los sistemas lineales son sensibles al error de redondeo y verá que cambios muy pequeños en los coeficientes pueden conducir a enormes imprecisiones en la solución. Por fortuna, hay algo que puede hacer para minimizar o incluso eliminar este problema (a diferencia del problema con los sistemas mal condicionados).

1. (a) Resuelva la ecuación lineal sencilla $0.00021x = 1$ para x.

   (b) Suponga que su calculadora puede acarrear sólo cuatro dígitos significativos. La ecuación se redondeará a $0.0002x = 1$. Resuelva esta ecuación.

La diferencia entre las respuestas de los incisos (a) y (b) puede considerarse como efecto de un error de 0.00001 en la solución de la ecuación dada.

2. Ahora extienda esta idea a un sistema de ecuaciones lineales.
   (a) Con eliminación gaussiana, resuelva el sistema lineal

$$0.400x + 99.6y = 100$$
$$75.3x - 45.3y = 30.0$$

   con tres dígitos significativos. Comience por pivotear en 0.400 y lleve cada cálculo a tres dígitos significativos. Debe obtener la "solución" $x = -1.00$, $y = 1.01$. Compruebe que la solución real es $x = 1.00$, $y = 1.00$. ¡Este es un error enorme, de -200% en el valor de x! ¿Puede descubrir qué lo causó?

   (b) Resuelva nuevamente el sistema del inciso (a), esta vez intercambiando las dos ecuaciones (o, de manera equivalente, los dos renglones de su matriz aumentada) y pivotee en 75.3. De nuevo, lleve cada cálculo a tres dígitos significativos. ¿Cuál es la solución esta vez?

La moraleja de la historia es que, cuando se usa eliminación gaussiana o de Gauss-Jordan para obtener una solución numérica de un sistema de ecuaciones lineales (esto es: una aproximación decimal), debe elegir el pivote con cuidado. Específicamente, en cada paso de pivoteo, elija de entre todos los pivotes posibles en una columna el elemento con el valor absoluto más grande. Use intercambios de renglones para llevar este elemento a la posición correcta y úselo para crear ceros donde se necesite en la columna. Esta estrategia se conoce como ***pivoteo parcial***.

3. Resuelva los siguientes sistemas mediante eliminación gaussiana, primero sin y luego con pivoteo parcial. Lleve cada cálculo a tres dígitos significativos. (Se proporcionan las soluciones exactas.)

(a) $$\begin{aligned} 0.001x + 0.995y &= 1.00 \\ -10.2x + 1.00y &= -50.0 \end{aligned}$$

Solución exacta: $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 5.00 \\ 1.00 \end{bmatrix}$

(b) $$\begin{aligned} 10x - 7y &= 7 \\ -3x + 2.09y + 6z &= 3.91 \\ 5x - y + 5z &= 6 \end{aligned}$$

Solución exacta: $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0.00 \\ -1.00 \\ 1.00 \end{bmatrix}$

Operaciones de conteo: introducción al análisis de algoritmos

Las eliminaciones gaussiana y de Gauss-Jordan son ejemplos de ***algoritmos***: procedimientos sistemáticos diseñados para implementar una tarea particular; en este caso, la reducción por renglones de la matriz aumentada de un sistema de ecuaciones lineales. Los algoritmos son particularmente adecuados para su implementación en computadoras, pero no todos los algoritmos se crean igual. Además de la rapidez, memoria y otros atributos del sistema de cómputo en el que corren, algunos algoritmos son más rápidos que otros. Una medida de la llamada *complejidad* de un algoritmo (una medida de su eficiencia o capacidad para realizar su tarea en un número razonable de pasos) es el número de operaciones básicas que realiza como función del número de variables que se ingresan.

Examine esta proposición en el caso de los dos algoritmos que se tienen para resolver un sistema lineal: eliminación gaussiana y de Gauss-Jordan. Para este propósito, las operaciones básicas son multiplicación y división; se supondrá que todas las otras operaciones se realizan con mucha más rapidez y pueden ignorarse. (Esta es una suposición razonable, mas no se intentará justificarla.) Considere sólo sistemas de ecuaciones con matrices de coeficientes *cuadradas*, de modo que, si la matriz de coeficientes es $n \times n$, el número de variables de entrada es n. Por ende, su tarea es encontrar el número de operaciones realizadas mediante eliminación gaussiana y de Gauss-Jordan como función de n. Más aún, no se preocupe por casos especiales que puedan surgir; en vez de ello, establezca el *peor caso* que pueda surgir: cuando el algoritmo tarda tanto como sea posible. Dado que esto proporcionará una estimación del tiempo que tardará una computadora en realizar el algoritmo (si se sabe cuánto tarda una computadora en realizar una sola operación), el número de operaciones realizadas por un algoritmo se denotará $T(n)$. Por lo general estará interesado en $T(n)$ para valores grandes de n, de modo que comparar esta función para diferentes algoritmos le permitirá determinar cuál tarda menos tiempo en ejecutarse.

Abu Ja'far Muhammad ibn Musa al-Khwarizmi (c. 780–850) fue un matemático persa cuyo libro *Hisab al-jabr w'al muqabalah* (c. 825) describió el uso de los numerales indoarábigos y las reglas de la aritmética básica. La segunda palabra del título del libro dio origen a la palabra *álgebra*, y la palabra *algoritmo* se deriva del nombre de al-Khwarizmi.

1. Considere la matriz aumentada

$$[A \mid \mathbf{b}] = \left[\begin{array}{ccc|c} 2 & 4 & 6 & 8 \\ 3 & 9 & 6 & 12 \\ -1 & 1 & -1 & 1 \end{array}\right]$$

Cuente el número de operaciones requeridas para llevar $[A \mid \mathbf{b}]$ a la forma escalonada por renglones

$$\left[\begin{array}{ccc|c} 1 & 2 & 3 & 4 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

(Por "operación" se entiende una multiplicación o una división.) Ahora cuente el número de operaciones necesarias para completar la fase de sustitución hacia atrás de la eliminación gaussiana. Registre el número total de operaciones.

2. Cuente el número de operaciones necesarias para realizar la eliminación de Gauss-Jordan; esto es: reducir $[A \mid \mathbf{b}]$ a su forma escalonada reducida por renglones

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

(donde se introducen ceros en cada columna inmediatamente después de crear el 1 pivote en dicha columna). ¿Qué sugieren sus respuestas acerca de la eficiencia relativa de los dos algoritmos?

Ahora se tratará de analizar los algoritmos en una forma general sistemática. Suponga que la matriz aumentada $[A \mid \mathbf{b}]$ surge de un sistema lineal con n ecuaciones y n variables; por tanto, $[A \mid \mathbf{b}]$ es $n \times (n+1)$:

$$[A \mid \mathbf{b}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right]$$

Suponga que los intercambios de renglón nunca se necesitan, que siempre puede crear un 1 pivote a partir de un pivote al dividir entre el pivote.

3. (a) Demuestre que se necesitan n operaciones para crear el primer 1 pivote:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & * & \cdots & * & * \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right]$$



(¿Por qué no necesita contar una operación para la creación del 1 pivote?) Ahora demuestre que se necesitan n operaciones para obtener el primer cero en la columna 1:

$$\left[\begin{array}{cccc|c} 1 & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array}\right]$$



(¿Por qué no necesita contar una operación para la creación del cero mismo?) Cuando la primera columna se ha “barrido”, se tiene la matriz

$$\left[\begin{array}{cccc|c} 1 & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & * & \cdots & * & * \end{array}\right]$$

Demuestre que el número total de operaciones necesarias hasta este punto es $n + (n - 1)n$.

(b) Demuestre que el número total de operaciones necesarias para alcanzar la forma escalonada por renglones

$$\left[\begin{array}{cccc|c} 1 & * & \cdots & * & * \\ 0 & 1 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & * \end{array}\right]$$

es

$$[n + (n - 1)n] + [(n - 1) + (n - 2)(n - 1)] + [(n - 2) + (n - 3)(n - 2)] + \cdots + [2 + 1 \cdot 2] + 1$$

que se simplifica a

$$n^2 + (n - 1)^2 + \cdots + 2^2 + 1^2$$

(c) Demuestre que el número de operaciones necesarias para completar la fase de sustitución hacia atrás es

$$1 + 2 + \cdots + (n - 1)$$

(d) Con las fórmulas de suma para las sumas en los incisos (b) y (c) (vea los ejercicios 51 y 52 de la sección 2.4 y el apéndice B), demuestre que el número total de operaciones, $T(n)$, realizadas mediante eliminación gaussiana es

$$T(n) = \tfrac{1}{3}n^3 + n^2 - \tfrac{1}{3}n$$

Dado que cada función polinominal está dominada por su término pivote para valores grandes de la variable, se ve que $T(n) \approx \frac{1}{3}n^3$ para grandes valores de n.

4. Demuestre que la eliminación de Gauss-Jordan tiene $T(n) \approx \frac{1}{2}n^3$ operaciones totales si se crean ceros arriba y abajo del 1 pivote conforme avanza. (Esto demuestra que, para sistemas grandes de ecuaciones lineales, la eliminación gaussiana es más rápida que esta versión de la eliminación de Gauss-Jordan.)

2.3 Conjuntos generadores e independencia lineal

El segundo de los tres caminos en el "trivium" trata con combinaciones lineales de vectores. Se vio que es posible ver la resolución de un sistema de ecuaciones lineales como preguntar si cierto vector es una combinación lineal de otros vectores. Esta idea se explora con más detalle en esta sección. Conduce a algunos conceptos muy importantes que se encontrarán repetidamente en capítulos posteriores.

Conjuntos generadores de vectores

Ahora se puede responder con facilidad la pregunta planteada en la sección 1.1: ¿cuándo un vector dado es una combinación lineal de otros vectores dados?

Ejemplo 2.18

(a) ¿El vector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ es una combinación lineal de los vectores $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}$?

(b) ¿$\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ es una combinación lineal de los vectores $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}$?

Solución

(a) Se quiere encontrar escalares x y y tales que

$$x\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + y\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Al desarrollar, se obtiene el sistema

$$\begin{aligned} x - y &= 1 \\ y &= 2 \\ 3x - 3y &= 3 \end{aligned}$$

cuya matriz aumentada es

$$\left[\begin{array}{rr|r} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 3 & -3 & 3 \end{array}\right]$$

(Observe que las columnas de la matriz aumentada son sólo los vectores dados; note el orden de los vectores, en particular, cuál vector es el vector constante.)

La forma escalonada reducida de esta matriz es

$$\left[\begin{array}{rr|r} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$



(Verifique esto.) De modo que la solución es $x = 3$, $y = 2$, y la correspondiente combinación lineal es

$$3\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} + 2\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

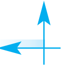
(b) Al usar la observación del inciso (a), se obtiene un sistema lineal cuya matriz aumentada es

$$\left[\begin{array}{rr|r} 1 & -1 & 2 \\ 0 & 1 & 3 \\ 3 & -3 & 4 \end{array}\right]$$

que se reduce a

$$\left[\begin{array}{rr|r} 1 & 0 & 5 \\ 0 & 1 & 3 \\ 0 & 0 & -2 \end{array}\right]$$

lo que revela que el sistema no tiene solución. Por tanto, en este caso, $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ no es una combinación lineal de $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}$.



La noción de conjunto generador está íntimamente conectada con la solución de sistemas lineales. Revise de nuevo el ejemplo 2.18. Ahí se vio que un sistema con matriz aumentada $[A \mid \mathbf{b}]$ tiene una solución precisamente cuando **b** es una combinación lineal de las columnas de A. Este es un hecho general que se resume en el siguiente teorema.

Teorema 2.4 Un sistema de ecuaciones lineales con matriz aumentada $[A \mid \mathbf{b}]$ es consistente si y sólo si **b** es una combinación lineal de las columnas de A.

Revise nuevamente el ejemplo 2.4 e interprételo a la luz del Teorema 2.4.

(a) El sistema

$$\begin{aligned} x - y &= 1 \\ x + y &= 3 \end{aligned}$$

tiene la solución única $x = 2, y = 1$. Por tanto,

$$2\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

Vea la figura 2.8(a).

(b) El sistema

$$\begin{aligned} x - y &= 2 \\ 2x - 2y &= 4 \end{aligned}$$

tiene un número infinito de soluciones de la forma $x = 2 + t, y = t$. Esto implica que

$$(2 + t)\begin{bmatrix} 1 \\ 2 \end{bmatrix} + t\begin{bmatrix} -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

para todos los valores de t. Geométricamente, los vectores $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\begin{bmatrix} -1 \\ -2 \end{bmatrix}$ y $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ son todos paralelos y por tanto todos yacen a lo largo de la misma recta que pasa por el origen [vea la figura 2.8(b)].

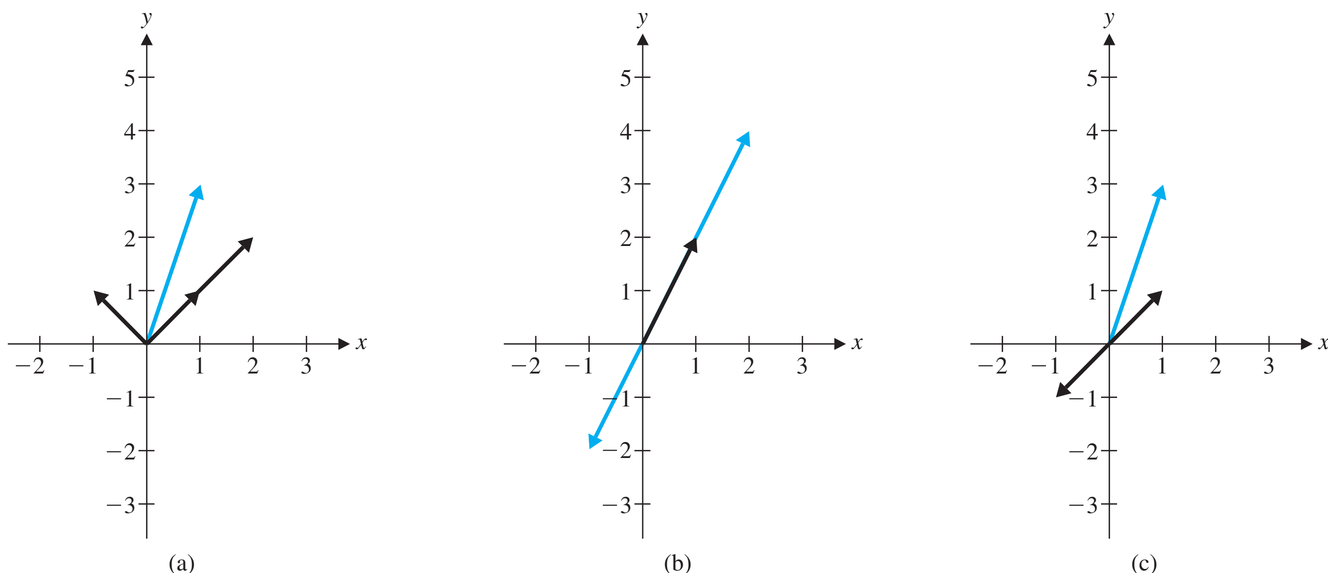


Figura 2.8

(c) El sistema

$$\begin{aligned} x - y &= 1 \\ x - y &= 3 \end{aligned}$$

no tiene soluciones, de modo que no hay valores de x y y que satisfagan

$$x\begin{bmatrix} 1 \\ 1 \end{bmatrix} + y\begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

En este caso, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ -1 \end{bmatrix}$ son paralelos, pero $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ no yace a lo largo de la misma recta que pasa por el origen [vea la figura 2.8(c)].

Con frecuencia se estará interesado en la colección de *todas* las combinaciones lineales de un conjunto dado de vectores.

Definición Si $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un conjunto de vectores en $\mathbb{R}^n$, entonces el conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ se llama el ***generador*** de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y se denota mediante gen $(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ o gen(S). Si gen $(S) = \mathbb{R}^n$, entonces S se llama ***conjunto generador*** de $\mathbb{R}^n$.

Ejemplo 2.19

Demuestre que $\mathbb{R}^2 = \text{gen}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$.

Solución Necesita demostrar que un vector arbitrario $\begin{bmatrix} a \\ b \end{bmatrix}$ puede escribirse como una combinación lineal de $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$; esto es, debe demostrar que la ecuación $x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$ siempre puede resolverse para x y y (en términos de a y b), sin importar los valores de a y b.

La matriz aumentada es $\left[\begin{array}{rr|r} 2 & 1 & a \\ -1 & 3 & b \end{array}\right]$ y la reducción por renglones produce

$$\left[\begin{array}{rr|r} 2 & 1 & a \\ -1 & 3 & b \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{rr|r} -1 & 3 & b \\ 2 & 1 & a \end{array}\right] \xrightarrow{R_2 + 2R_1} \left[\begin{array}{rr|c} -1 & 3 & b \\ 0 & 7 & a + 2b \end{array}\right]$$

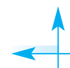
en cuyo punto es claro que el sistema tiene una solución (única). (¿Por qué?) Si continúa, obtiene

$$\xrightarrow{\frac{1}{7}R_2} \left[\begin{array}{rr|c} -1 & 3 & b \\ 0 & 1 & (a + 2b)/7 \end{array}\right] \xrightarrow{R_1 - 3R_2} \left[\begin{array}{rr|c} -1 & 0 & (b - 3a)/7 \\ 0 & 1 & (a + 2b)/7 \end{array}\right]$$

de donde se ve que $x = (3a - b)/7$ y $y = (a + 2b)/7$. Por tanto, para cualquier elección de a y b, se tiene

$$\left(\frac{3a - b}{7}\right)\begin{bmatrix} 2 \\ -1 \end{bmatrix} + \left(\frac{a + 2b}{7}\right)\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$


(Compruebe esto.)



Comentario También es cierto que $\mathbb{R}^2 = \text{gen}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 7 \end{bmatrix}\right)$: si, dado $\begin{bmatrix} a \\ b \end{bmatrix}$, puede encontrar x y y tales que $x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$, entonces también se tiene $x\begin{bmatrix} 2 \\ -1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 3 \end{bmatrix} + 0\begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$. De hecho, cualquier conjunto de vectores que *contenga* un conjunto generador para $\mathbb{R}^2$ también será un conjunto generador para $\mathbb{R}^2$ (vea el ejercicio 20).

El siguiente ejemplo es un importante caso (sencillo) de un conjunto generador. Muchas veces encontrará versiones de este ejemplo.

Ejemplo 2.20

Sean $\mathbf{e}_1$, $\mathbf{e}_2$, y $\mathbf{e}_3$ los vectores unitarios estándar en $\mathbb{R}^3$. Entonces, para cualquier vector $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$, se tiene

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3$$

Por tanto, $\mathbb{R}^3 = \text{gen}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

No debe tener dificultad para ver que, en general, $\mathbb{R}^n = \text{gen}(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$.

Cuando el generador de un conjunto de vectores en $\mathbb{R}^n$ no es todo de $\mathbb{R}^n$, es razonable pedir una descripción del generador de los vectores.

Ejemplo 2.21

Encuentre el generador de $\begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \\ -3 \end{bmatrix}$. (Vea el ejemplo 2.18.)

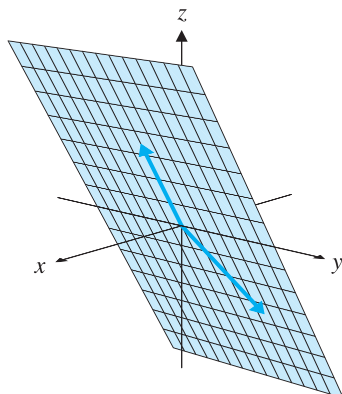


Figura 2.9
Dos vectores no paralelos generan un plano

Solución Al pensar geométricamente, puede ver que el conjunto de todas las combinaciones lineales de $\begin{bmatrix}1\\0\\3\end{bmatrix}$ y $\begin{bmatrix}-1\\1\\-3\end{bmatrix}$ es justo el plano que pasa por el origen con $\begin{bmatrix}1\\0\\3\end{bmatrix}$ y $\begin{bmatrix}-1\\1\\-3\end{bmatrix}$ como vectores directores (figura 2.9). La ecuación vectorial de este plano es $\begin{bmatrix}x\\y\\z\end{bmatrix} = s\begin{bmatrix}1\\0\\3\end{bmatrix} + t\begin{bmatrix}-1\\1\\-3\end{bmatrix}$, que es sólo otra forma de decir que $\begin{bmatrix}x\\y\\z\end{bmatrix}$ está en el generador de $\begin{bmatrix}1\\0\\3\end{bmatrix}$ y $\begin{bmatrix}-1\\1\\-3\end{bmatrix}$.

Suponga que quiere obtener la ecuación general de este plano. Existen muchas formas de proceder. Una es usar el hecho de que la ecuación $ax + by + cz = 0$ debe ser satisfecha por los puntos $(1, 0, 3)$ y $(-1, 1, -3)$ determinados por los vectores directores. Entonces la sustitución conduce a un sistema de ecuaciones en a, b y c. (Vea el ejercicio 17.)

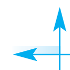
Otro método es usar el sistema de ecuaciones que surgen de la ecuación vectorial:

$$\begin{aligned} s - t &= x \\ t &= y \\ 3s - 3t &= z \end{aligned}$$

Si la matriz aumentada se reduce por renglones, se obtiene

$$\left[\begin{array}{rr|c} 1 & -1 & x \\ 0 & 1 & y \\ 3 & -3 & z \end{array}\right] \xrightarrow{R_3 - 3R_1} \left[\begin{array}{rr|c} 1 & -1 & x \\ 0 & 1 & y \\ 0 & 0 & z - 3x \end{array}\right]$$

Ahora se sabe que este sistema es inconsistente, pues $\begin{bmatrix}x\\y\\z\end{bmatrix}$ *está* en el generador de $\begin{bmatrix}1\\0\\3\end{bmatrix}$ y $\begin{bmatrix}-1\\1\\-3\end{bmatrix}$ por suposición. De modo que se *debe* tener $z - 3x = 0$, (o $3x - z = 0$, en la forma más estándar), lo que da la ecuación general que se busca.



Comentario Un vector normal al plano en este ejemplo también está dado por el producto cruz

$$\begin{bmatrix}1\\0\\3\end{bmatrix} \times \begin{bmatrix}-1\\1\\-3\end{bmatrix} = \begin{bmatrix}-3\\0\\1\end{bmatrix}$$

Independencia lineal

En el ejemplo 2.18 se encontró que $3\begin{bmatrix}1\\0\\3\end{bmatrix} + 2\begin{bmatrix}-1\\1\\-3\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix}$. Abrevie esta ecuación como $3\mathbf{u} + 2\mathbf{v} = \mathbf{w}$. El vector $\mathbf{w}$ "depende" de $\mathbf{u}$ y $\mathbf{v}$ en el sentido de que es una combinación lineal de ellos. Se dice que un conjunto de vectores es ***linealmente dependiente*** si uno de ellos puede escribirse como una combinación lineal de los otros. Note que tam-

bién se tiene $\mathbf{u} = -\frac{2}{3}\mathbf{v} + \frac{1}{3}\mathbf{w}$ y $\mathbf{v} = -\frac{3}{2}\mathbf{u} + \frac{1}{2}\mathbf{w}$. Para franquear la cuestión de cuál vector expresar en términos del resto, a continuación se enuncia la definición formal:

Definición Un conjunto de vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ es ***linealmente dependiente*** si existen escalares $c_1, c_2, \ldots, c_k$, *al menos uno de los cuales no es cero*, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

Un conjunto de vectores que no es linealmente dependiente se llama ***linealmente independiente***.

Comentarios

- En la definición de independencia lineal, el requisito de que al menos uno de los escalares $c_1, c_2, \ldots, c_k$ debe ser distinto de cero admite la posibilidad de que alguno puede ser cero. En el ejemplo anterior, $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ son linealmente dependientes, pues $3\mathbf{u} + 2\mathbf{v} - \mathbf{w} = \mathbf{0}$ y, de hecho, *todos* los escalares son distintos de cero. Por otra parte,

$$\begin{bmatrix} 2 \\ 6 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 3 \end{bmatrix} + 0\begin{bmatrix} 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

de modo que $\begin{bmatrix} 2 \\ 6 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ son linealmente dependientes, pues al menos uno (de hecho, dos) de los tres escalares 1, -2 y 0 es distinto de cero. (Note que la dependencia real surge simplemente del hecho de que los primeros dos vectores son múltiplos.) (Vea el ejercicio 44.)

- Dado que $0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k = \mathbf{0}$ para *cualesquiera* vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$, la dependencia lineal en esencia dice que el vector cero puede expresarse como una combinación lineal *no trivial* de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Por tanto, independencia lineal significa que el vector cero puede expresarse como una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ *sólo* en la forma trivial: $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$ sólo si $c_1 = 0, c_2 = 0, \ldots, c_k = 0$.

La relación entre la noción intuitiva de dependencia y la definición formal se proporciona en el siguiente teorema. ¡Felizmente, las dos nociones son equivalentes!

Teorema 2.5 Los vectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ en $\mathbb{R}^n$ son linealmente dependientes si y sólo si al menos uno de los vectores puede expresarse como una combinación lineal de los otros.

Demostración Si uno de los vectores, por ejemplo, $\mathbf{v}_1$ es una combinación lineal de los otros, entonces existen escalares $c_2, \ldots, c_m$ tales que $\mathbf{v}_1 = c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m$. Al reordenar, se obtiene $\mathbf{v}_1 - c_2\mathbf{v}_2 - \cdots - c_m\mathbf{v}_m = \mathbf{0}$, lo que implica que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes, pues al menos uno de los escalares (a saber, el coeficiente 1 de $\mathbf{v}_1$) es distinto de cero.

Por el contrario, suponga que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes. Entonces existen escalares $c_1, c_2, \ldots, c_m$, no todos cero, tales que $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$. Suponga $c_1 \neq 0$. Entonces

$$c_1\mathbf{v}_1 = -c_2\mathbf{v}_2 - \cdots - c_m\mathbf{v}_m$$

y se pueden multiplicar ambos lados por $1/c_1$ para obtener $\mathbf{v}_1$ como una combinación lineal de los otros vectores:

$$\mathbf{v}_1 = -\left(\frac{c_2}{c_1}\right)\mathbf{v}_2 - \cdots - \left(\frac{c_m}{c_1}\right)\mathbf{v}_m$$

Nota Parece como si se hiciera un poco de trampa en esta demostración. Después de todo, no puede estar seguro de que $\mathbf{v}_1$ es una combinación lineal de los otros vectores, ni que c_1 es distinto de cero. Sin embargo, el argumento es análogo para algún otro vector $\mathbf{v}_i$ o para un escalar c_j diferente. Alternativamente, puede reetiquetar las cosas de modo que funcionen como en la demostración anterior. En una situación como ésta, un matemático puede comenzar por decir, "sin pérdida de generalidad, puede suponer que $\mathbf{v}_1$ es una combinación lineal de los otros vectores" y luego proceder como antes.

Ejemplo 2.22

Cualquier conjunto de vectores que contenga el vector cero es linealmente dependiente. Pues si $\mathbf{0}, \mathbf{v}_2, \ldots, \mathbf{v}_m$ están en $\mathbb{R}^n$, entonces puede encontrar una combinación no trivial de la forma $c_1\mathbf{0} + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$ al hacer $c_1 = 1$ y $c_2 = c_3 = \cdots = c_m = 0$.

Ejemplo 2.23

Determine si los siguientes conjuntos de vectores son linealmente independientes:

(a) $\begin{bmatrix} 1 \\ 4 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$

(b) $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

(c) $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$

Solución Al responder cualquier pregunta de este tipo, es buena idea ver si puede determinar por inspección si un vector es una combinación lineal de los otros. ¡Pensar un poco puede ahorrar muchos cálculos!

(a) La única forma en que los dos vectores pueden ser linealmente dependientes es si uno es múltiplo del otro. (¿Por qué?) Claramente, estos dos vectores no son múltiplos, de modo que son linealmente independientes.

(b) No hay aquí una relación de dependencia obvia, así que se tratará de encontrar escalares c_1, c_2, c_3 tales que

$$c_1\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + c_3\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

El sistema lineal correspondiente es

$$\begin{aligned} c_1 + \phantom{c_2 + {}} c_3 &= 0 \\ c_1 + c_2 \phantom{{} + c_3} &= 0 \\ c_2 + c_3 &= 0 \end{aligned}$$

y la matriz aumentada es

$$\left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right]$$

Una vez más, se hace la observación fundamental de que ¡las columnas de la matriz de coeficientes son justo los vectores en cuestión!

La forma escalonada reducida por renglones es

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$



(compruebe esto), de modo que $c_1 = 0, c_2 = 0, c_3 = 0$. Por tanto, los vectores dados son linealmente independientes.

(c) Un poco de reflexión revela que

$$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$



de modo que los tres vectores son linealmente dependientes. [Establezca un sistema lineal como en el inciso (b) para comprobar esto algebraicamente.]

(d) Una vez más, no se observa dependencia obvia, así que se procede directamente a reducir un sistema lineal homogéneo cuya matriz aumentada tiene como sus columnas los vectores dados:

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 2 & 1 & 4 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right] \xrightarrow{R_2 - 2R_1} \left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & -1 & 2 & 0 \\ 0 & -1 & 2 & 0 \end{array}\right] \xrightarrow[\substack{R_3 - R_2 \\ -R_2}]{R_1 + R_2} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Si se hace que los escalares sean c_1, c_2 y c_3, se tiene

$$\begin{aligned} c_1 \quad + 3c_3 &= 0 \\ c_2 - 2c_3 &= 0 \end{aligned}$$

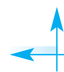
de donde se ve que el sistema tiene un número infinito de soluciones. En particular, debe haber una solución distinta de cero, de modo que los vectores dados son linealmente dependientes.

Si continúa, puede describir estas soluciones exactamente: $c_1 = -3c_3$ y $c_2 = 2c_3$. Por tanto, para cualquier valor distinto de cero de c_3, se tiene la relación de dependencia lineal

$$-3c_3\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + 2c_3\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} + c_3\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$



(Una vez más, compruebe que esto es correcto.)



Este procedimiento se resume para probar la independencia lineal como teorema.

Teorema 2.6 Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ vectores (columna) en $\mathbb{R}^n$ y sea A la matriz de $n \times m$ $[\mathbf{v}_1\ \mathbf{v}_2 \cdots \mathbf{v}_m]$ con dichos vectores como sus columnas. Entonces $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes si y sólo si el sistema lineal homogéneo con matriz aumentada $[A \mid \mathbf{0}]$ tiene una solución no trivial.

Demostración $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes si y sólo si existen escalares $c_1, c_2, \ldots, c_m$, no todos cero, tales que $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m = \mathbf{0}$. Por el Teorema 2.4, esto es equivalente a decir que el vector distinto de cero $\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$ es una solución del sistema cuya matriz aumentada es $[\mathbf{v}_1\ \mathbf{v}_2 \cdots \mathbf{v}_m \mid \mathbf{0}]$.

Ejemplo 2.24

Los vectores unitarios estándar $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ son linealmente dependientes en $\mathbb{R}^3$, pues el sistema con matriz aumentada $[\mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3 \mid \mathbf{0}]$ ya está en la forma escalonada reducida por renglones

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

y por tanto claramente sólo tiene la solución trivial. En general, se ve que $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ serán linealmente independientes en $\mathbb{R}^n$.

Al realizar operaciones elementales con renglones sobre una matriz se construyen combinaciones lineales de los renglones. Puede usar este hecho para encontrar otra forma de probar la independencia lineal para vectores.

Ejemplo 2.25

Considere los tres vectores del ejemplo 2.23(d) como vectores *renglón*:

$$[1, 2, 0], \quad [1, 1, -1] \quad \text{y} \quad [1, 4, 2]$$

Construya una matriz con estos vectores como sus renglones y proceda a reducirla a forma escalonada. Cada vez que cambia un renglón, el nuevo renglón se denota al agregar un símbolo de prima:

$$\begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & -1 \\ 1 & 4 & 2 \end{bmatrix} \xrightarrow[R_3' = R_3 - R_1]{R_2' = R_2 - R_1} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & -1 \\ 0 & 2 & 2 \end{bmatrix} \xrightarrow{R_3'' = R_3' + 2R_2'} \begin{bmatrix} 1 & 2 & 0 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

A partir de esto se ve que

$$\mathbf{0} = R_3'' = R_3' + 2R_2' = (R_3 - R_1) + 2(R_2 - R_1) = -3R_1 + 2R_2 + R_3$$

o, en términos de los vectores originales,

$$-3[1, 2, 0] + 2[1, 1, -1] + [1, 4, 2] = [0, 0, 0]$$

[Note que este planteamiento corresponde a tomar $c_3 = 1$ en la solución del ejemplo 2.23(d).]

Por tanto, los renglones de una matriz serán linealmente dependientes si pueden usarse operaciones elementales con renglones para crear un renglón cero. Este hallazgo se resume del modo siguiente:

Teorema 2.7

Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ vectores (renglón) en $\mathbb{R}^n$ y sea A la matriz de $m \times n$ $\begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_m \end{bmatrix}$ con dichos vectores como sus renglones. Entonces $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes si y sólo si rango$(A) < m$.

Demostración Suponga que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes. Entonces, por el Teorema 2.2, al menos uno de los vectores puede escribirse como una combinación lineal de los otros. Reetiquete los vectores, si es necesario, de modo que pueda escribir $\mathbf{v}_m$

$= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{m-1}\mathbf{v}_{m-1}$. Entonces las operaciones elementales con renglones $R_m - c_1R_1, R_m - c_2R_2, \ldots, R_m - c_{m-1}R_{m-1}$ aplicadas a A crearán un renglón cero en el renglón m. Por tanto, rango$(A) < m$.

Por el contrario, suponga que rango$(A) < m$. Entonces existe alguna secuencia de operaciones con renglones que crearán un renglón cero. Puede usar un argumento de sustitución sucesiva análogo al que se utilizó en el ejemplo 2.25 para demostrar que $\mathbf{0}$ es una combinación lineal no trivial de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. Por tanto, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes.

En algunas situaciones, puede deducir que un conjunto de vectores es linealmente dependiente sin hacer trabajo alguno. Una de tales situaciones es cuando el vector cero está en el conjunto (como en el ejemplo 2.22). Otra es cuando hay "demasiados" vectores para que sean independientes. El siguiente teorema resume este caso. (En el capítulo 6 se verá una versión más precisa de este resultado.)

Teorema 2.8 Cualquier conjunto de m vectores en $\mathbb{R}^n$ es linealmente dependiente si $m > n$.

Demostración Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ vectores (columna) en $\mathbb{R}^n$ y sea A la matriz $n \times m$ $[\mathbf{v}_1\ \mathbf{v}_2 \cdots \mathbf{v}_m]$ con dichos vectores como sus columnas. Por el Teorema 2.6, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes si y sólo si el sistema lineal homogéneo con matriz aumentada $[A \mid \mathbf{0}]$ tiene una solución no trivial. Pero, de acuerdo con el Teorema 2.6, este siempre será el caso si A tiene más columnas que renglones; es el caso aquí, pues el número de columnas m es mayor que el número de renglones n.

Ejemplo 2.26

Los vectores $\begin{bmatrix}1\\3\end{bmatrix}, \begin{bmatrix}2\\4\end{bmatrix}$ y $\begin{bmatrix}3\\1\end{bmatrix}$ son linealmente dependientes, pues no pueden existir más de dos vectores linealmente independientes en $\mathbb{R}^2$. (Note que si quiere encontrar la relación de dependencia real entre estos tres vectores, debe resolver el sistema homogéneo cuya matriz de coeficientes tenga los vectores dados como columnas. ¡Hágalo!)

Ejercicios 2.3

En los ejercicios 1-6, determine si el vector $\mathbf{v}$ *es una combinación lineal de los vectores restantes.*

1. $\mathbf{v} = \begin{bmatrix}1\\2\end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix}1\\-1\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}2\\-1\end{bmatrix}$

2. $\mathbf{v} = \begin{bmatrix}1\\2\end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix}-1\\3\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}2\\-6\end{bmatrix}$

3. $\mathbf{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}0\\1\\1\end{bmatrix}$

4. $\mathbf{v} = \begin{bmatrix}3\\1\\-2\end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}0\\1\\1\end{bmatrix}$

5. $\mathbf{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix}1\\1\\0\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}0\\1\\1\end{bmatrix},$

$\mathbf{u}_3 = \begin{bmatrix}1\\0\\1\end{bmatrix}$

CAS **6.** $\mathbf{v} = \begin{bmatrix}2.2\\4.0\\-2.2\end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix}1.0\\0.5\\-0.5\end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix}2.4\\1.2\\3.1\end{bmatrix},$

$\mathbf{u}_3 = \begin{bmatrix}1.2\\-2.3\\4.8\end{bmatrix}$

En los ejercicios 7 y 8, determine si el vector **b** *está en el generador de las columnas de la matriz A.*

7. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

8. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 10 \\ 11 \\ 12 \end{bmatrix}$

9. Demuestre que $\mathbb{R}^2 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$.

10. Demuestre que $\mathbb{R}^2 = \text{gen}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}\right)$.

11. Demuestre que $\mathbb{R}^3 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right)$.

12. Demuestre que $\mathbb{R}^3 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\right)$.

En los ejercicios 13-16, describa el generador de los vectores dados (a) geométricamente y (b) algebraicamente.

13. $\begin{bmatrix} 2 \\ -4 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \end{bmatrix}$

14. $\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

15. $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$

16. $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

17. La ecuación general del plano que contiene los puntos $(1, 0, 3)$, $(-1, 1, -3)$ y el origen es de la forma $ax + by + cz = 0$. Resuelva para a, b y c.

18. Demuestre que $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ están todos en $\text{gen}(\mathbf{u}, \mathbf{v}, \mathbf{w})$.

19. Demuestre que $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ están todos en $\text{gen}(\mathbf{u}, \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} + \mathbf{w})$.

20. **(a)** Demuestre que si $\mathbf{u}_1, \ldots, \mathbf{u}_m$ son vectores en $\mathbb{R}^n$, $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k\}$, y $T = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{u}_{k+1}, \ldots, \mathbf{u}_m\}$, entonces; $\text{gen}(S) \subseteq \text{gen}(T)$. [*Sugerencia:* vuelva a escribir esta ecuación en términos de combinaciones lineales.]

(b) Deduzca que, si $\mathbb{R}^n = \text{gen}(S)$, entonces $\mathbb{R}^n = \text{gen}(T)$ también.

21. **(a)** Suponga que el vector $\mathbf{w}$ es una combinación lineal de los vectores $\mathbf{u}_1, \ldots, \mathbf{u}_k$ y que cada $\mathbf{u}_i$ es una combinación lineal de los vectores $\mathbf{v}_1, \ldots, \mathbf{v}_m$. Demuestre que $\mathbf{w}$ es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_m$ y por tanto $\text{gen}(\mathbf{u}_1, \ldots, \mathbf{u}_k) \subseteq \text{gen}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$.

(b) En el inciso (a), suponga además que cada $\mathbf{v}_j$ también es una combinación lineal de $\mathbf{u}_1, \ldots, \mathbf{u}_k$. Demuestre que $\mathbf{u}_1, \ldots, \mathbf{u}_k) = \text{gen}(\mathbf{v}_1, \ldots, \mathbf{v}_m)$.

(c) Use el resultado del inciso (b) para demostrar que

$$\mathbb{R}^3 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$$

[*Sugerencia:* Se sabe que $\mathbb{R}^3 = \text{gen}(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.]

Use el método del ejemplo 2.23 y el Teorema 2.6 para determinar si los conjuntos de vectores en los ejercicios 22-31 son linealmente independientes. Si, para alguno de ellos la respuesta puede determinarse por inspección (esto es, sin cálculo), establezca por qué. Para cualquier conjunto que sea linealmente dependiente, encuentre una relación de dependencia entre los vectores.

22. $\begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$

23. $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

24. $\begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -6 \end{bmatrix}$

25. $\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$

26. $\begin{bmatrix} 2 \\ -3 \\ 7 \end{bmatrix}, \begin{bmatrix} -5 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}$

27. $\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \\ 8 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

28. $\begin{bmatrix} 1 \\ 4 \\ 3 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ 2 \\ 4 \end{bmatrix}$

29. $\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \end{bmatrix}$

30. $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \\ 2 \\ 1 \end{bmatrix}$

31. $\begin{bmatrix} 3 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 3 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 3 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \\ 3 \end{bmatrix}$

En los ejercicios 32-41, determine si los conjuntos de vectores en el ejercicio dado son linealmente independientes al convertir los vectores a vectores renglón y usar el método del ejemplo

2.25 y el Teorema 2.7. Para cualquier conjunto que sea linealmente independiente, encuentre una relación de dependencia entre los vectores.

32. Ejercicio 22 **33.** Ejercicio 23

34. Ejercicio 24 **35.** Ejercicio 25

36. Ejercicio 26 **37.** Ejercicio 27

38. Ejercicio 28 **39.** Ejercicio 29

40. Ejercicio 30 **41.** Ejercicio 31

42. **(a)** Si las columnas de una matriz A de $n \times n$ son linealmente independientes como vectores en $\mathbb{R}^n$, ¿cuál es el rank de A? Explique.

(b) Si los renglones de una matriz A de $n \times n$ son linealmente independientes como vectores en $\mathbb{R}^n$, ¿cuál es el rank de A? Explique.

43. **(a)** Si los vectores **u**, **v** y **w** son linealmente independientes, ¿$\mathbf{u} + \mathbf{v}$, $\mathbf{v} + \mathbf{w}$ y $\mathbf{u} + \mathbf{w}$ también son linealmente independientes? Justifique su respuesta.

(b) Si los vectores **u**, **v** y **w** son linealmente independientes, ¿$\mathbf{u} - \mathbf{v}$, $\mathbf{v} - \mathbf{w}$ y $\mathbf{u} - \mathbf{w}$ también son linealmente independientes? Justifique su respuesta.

44. Demuestre que dos vectores son linealmente dependientes si y sólo si uno es un múltiplo escalar del otro. [*Sugerencia:* considere por separado el caso en que uno de los vectores es **0**.]

45. Ofrezca una "demostración de vector renglón" del Teorema 2.8.

46. Demuestre que todo subconjunto de un conjunto linalmente independiente es linealmente independiente.

47. Suponga que $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}\}$ es un conjunto de vectores en algún $\mathbb{R}^n$ y que **v** es una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_k$.. Si $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$, demuestre que gen$(S)$ = gen(S'). [*Sugerencia:* el ejercicio 21(b) es útil aquí.]

48. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto linealmente independiente de vectores en $\mathbb{R}^n$ y sea **v** un vector en $\mathbb{R}^n$. Suponga que $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ con $c_1 \neq 0$. Pruebe que $\{\mathbf{v}, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es linealmente independiente.

2.4 Aplicaciones

Existen demasiadas aplicaciones de los sistemas de ecuaciones lineales para hacerles justicia en una sola sección. Esta sección presentará algunas aplicaciones para ilustrar la diversidad de escenarios en que surgen.

Asignación de recursos

Una gran cantidad de aplicaciones de los sistemas de ecuaciones lineales involucran asignar recursos limitados sujetos a un conjunto de restricciones.

Ejemplo 2.27

Una bióloga colocó tres cepas de bacterias (denominadas I, II y III) en un tubo de ensayo, donde se alimentarán de tres diferentes fuentes alimenticias (A, B y C). Cada día, 2300 unidades de A, 800 unidades de B y 1500 unidades de C se colocan en el tubo de ensayo y cada bacteria consume cierto número de unidades de cada alimento por día, como se muestra en la tabla 2.2. ¿Cuántas bacterias de cada cepa pueden coexistir en el tubo de ensayo y consumir todo el alimento?

Tabla 2.2

	Bacteria cepa I	Bacteria cepa II	Bacteria cepa III
Alimento A	2	2	4
Alimento B	1	2	0
Alimento C	1	3	1

Solución Sean x_1, x_2 y x_3 los números de bacterias de las cepas I, II y III, respectivamente. Dado que cada una de las bacterias x_1 de la cepa I consumen 2 unidades de A por día, la cepa I consume un total de $2x_1$ unidades por día. De igual modo, las cepas II y III consumen un total de $2x_2$ y $4x_3$ unidades de alimento A diariamente. Puesto que se quiere consumir las 2300 unidades de A, se tiene la ecuación

$$2x_1 + 2x_2 + 4x_3 = 2300$$

Del mismo modo, se obtienen ecuaciones correspondientes para el consumo de B y C:

$$\begin{aligned} x_1 + 2x_2 \phantom{{}+x_3} &= 800 \\ x_1 + 3x_2 + x_3 &= 1500 \end{aligned}$$

Por tanto, se tiene un sistema de tres ecuaciones lineales con tres variables. La reducción por renglones de la matriz aumentada correspondiente produce

$$\left[\begin{array}{ccc|c} 2 & 2 & 4 & 2300 \\ 1 & 2 & 0 & 800 \\ 1 & 3 & 1 & 1500 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 100 \\ 0 & 1 & 0 & 350 \\ 0 & 0 & 1 & 350 \end{array}\right]$$

En consecuencia, $x_1 = 100$, $x_2 = 350$ y $x_3 = 350$. La bióloga debe colocar 100 bacterias de cepa I y 350 de cada una de las cepas II y II en su tubo de ensayo si quiere que consuman todo el alimento.

Ejemplo 2.28

Repita el ejemplo 2.27 y use los datos de consumo diario de alimento (unidades por día) que se muestran en la tabla 2.3. Suponga esta vez que en el tubo de ensayo se colocan diariamente 1500 unidades de A, 3000 unidades de B y 4500 unidades de C.

Tabla 2.3

	Bacteria cepa I	Bacteria cepa II	Bacteria cepa III
Alimento A	1	1	1
Alimento B	1	2	3
Alimento C	1	3	5

Solución Sean nuevamente x_1, x_2 y x_3 los números de bacterias de cada tipo. La matriz aumentada para el sistema lineal resultante y la correspondiente forma escalonada reducida son

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1500 \\ 1 & 2 & 3 & 3000 \\ 1 & 3 & 5 & 4500 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 1500 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Se ve que, en este caso, se tiene más de una solución, dadas por

$$\begin{aligned} x_1 \phantom{{}+2x_2} - x_3 &= 0 \\ x_2 + 2x_3 &= 1500 \end{aligned}$$

Al hacer $x_3 = t$, se obtiene $x_1 = t$, $x_2 = 1500 - 2t$ y $x_3 = t$. En cualquier problema de aplicación, debe tener cuidado de interpretar adecuadamente las soluciones. Ciertamente el número de bacterias no puede ser negativo. Por tanto, $t \geq 0$ y $1500 - 2t \geq 0$. La última

desigualdad implica que $t \le 750$, de modo que se tiene $0 \le t \le 750$. Presumiblemente, el número de bacterias debe ser un número entero, de modo que existen exactamente 751 valores de t que satisfacen la desigualdad. Por ende, las 751 soluciones son de la forma

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} t \\ 1500 - 2t \\ t \end{bmatrix} = \begin{bmatrix} 0 \\ 1500 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

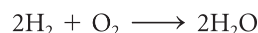
una para cada valor entero de t tal que $0 \le t \le 750$. (De este modo, aunque matemáticamente este sistema tiene un número infinito de soluciones, *físicamente* sólo existe un número finito.)



Balanceo de ecuaciones químicas

Cuando ocurre una reacción química, ciertas moléculas (los *reactivos*) se combinan para formar nuevas moléculas (los *productos*). Una ***ecuación química balanceada*** es una ecuación algebraica que proporciona los números relativos de reactivos y productos en la reacción, y tiene el mismo número de átomos de cada tipo en los lados izquierdo y derecho. La ecuación usualmente se escribe con los reactivos a la izquierda, los productos a la derecha y una flecha entre ellos para mostrar la dirección de la reacción.

Por ejemplo, para la reacción en la que se combinan gas hidrógeno (H_2) y oxígeno (O_2) para formar agua (H_2O), una ecuación química balanceada es



$$2H_2 + O_2 \longrightarrow 2H_2O$$

que indica que dos moléculas de hidrógeno se combinan con una molécula de oxígeno para formar dos moléculas de agua. Observe que la ecuación está balanceada, pues existen cuatro átomos de hidrógeno y dos átomos de oxígeno en cada lado. Note que nunca habrá una sola ecuación balanceada para una reacción, pues cualquier múltiplo entero positivo de una ecuación balanceada también lo estará. Por ejemplo, $6H_2 + 3O_2 \longrightarrow 6H_2O$ también está balanceada. Por tanto, por lo general se busca la ecuación balanceada *más simple* para una reacción dada.

Aunque frecuentemente ensayo y error funcionan en los ejemplos simples, el proceso de balancear ecuaciones químicas en realidad involucra resolver un sistema homogéneo de ecuaciones lineales, de modo que pueden usarse las técnicas desarrolladas para remover las conjeturas.

Ejemplo 2.29

La combustión de amoniaco (NH_3) en oxígeno produce nitrógeno (N_2) y agua. Encuentre una ecuación química balanceada para esta reacción.

Solución Si el número de moléculas de amoniaco, oxígeno, nitrógeno y agua se denotan w, x, y y z, respectivamente, entonces se busca una ecuación de la forma



$$wNH_3 + xO_2 \longrightarrow yN_2 + zH_2O$$

Al comparar el número de átomos de nitrógeno, hidrógeno y oxígeno en los reactivos y productos, se obtienen tres ecuaciones lineales:

$$\begin{aligned} &\text{Nitrógeno:} && w = 2y \\ &\text{Hidrógeno:} && 3w = 2z \\ &\text{Oxígeno:} && 2x = z \end{aligned}$$

Al reescribir estas ecuaciones en forma estándar se obtiene un sistema homogéneo de tres ecuaciones lineales con cuatro variables. [Note que el Teorema 2.3 garantiza que tal

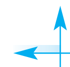
sistema tendrá (un número infinito de) soluciones no triviales.] Reduzca la correspondiente matriz aumentada mediante eliminación de Gauss-Jordan.

$$\begin{array}{rcrcrcrcl} w & & & - & 2y & & & = & 0 \\ 3w & & & & & - & 2z & = & 0 \\ & & 2x & & & - & z & = & 0 \end{array} \longrightarrow \left[\begin{array}{rrrr|r} 1 & 0 & -2 & 0 & 0 \\ 3 & 0 & 0 & -2 & 0 \\ 0 & 2 & 0 & -1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrrr|r} 1 & 0 & 0 & -\frac{2}{3} & 0 \\ 0 & 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & -\frac{1}{3} & 0 \end{array}\right]$$

Por tanto, $w = \frac{2}{3}z$, $x = \frac{1}{2}z$ y $y = \frac{1}{3}z$. El valor positivo más pequeño de z que producirá valores *enteros* para las cuatro variables es el mínimo común denominador de las fracciones $\frac{2}{3}$, $\frac{1}{2}$ y $\frac{1}{3}$, a saber, 6, lo que produce $w = 4$, $x = 3$, $y = 2$ y $z = 6$. Por tanto, la ecuación química balanceada es



$$4NH_3 + 3O_2 \longrightarrow 2N_2 + 6H_2O$$



Análisis de redes

Muchas situaciones prácticas originan redes: redes de transporte, redes de comunicaciones y redes económicas, por nombrar algunas. De particular interés son los posibles *flujos* a través de las redes. Por ejemplo, los vehículos que fluyen a través de una red de carreteras, la información que fluye a través de una red de datos, y los bienes y servicios que fluyen a través de una red económica.

En este texto, una ***red*** consistirá de un número finito de ***nodos*** (también llamados ***uniones*** o ***vértices***) conectados mediante una serie de aristas dirigidas llamadas ***ramas*** o ***arcos***. Cada rama se marcará con un ***flujo*** que representa la cantidad de algún objeto que puede fluir a lo largo o a través de cada rama en la dirección indicada. (Piense en los automóviles que viajan a lo largo de una red de calles de un sentido.) La regla fundamental que gobierna el flujo a través de una red es la ***conservación del flujo***:

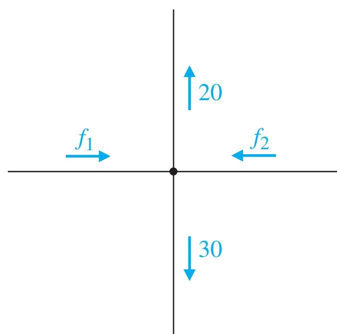


Figura 2.10
Flujo en un nodo: $f_1 + f_2 = 50$

> En cada nodo, el flujo de entrada es igual al flujo de salida.

La figura 2.10 muestra una porción de una red, con dos ramas que entran a un nodo y dos que salen. La regla de conservación del flujo implica que el flujo entrante total, $f_1 + f_2$ unidades, debe coincidir con el flujo saliente total, 20 + 30 unidades. Por tanto, se tiene la ecuación lineal $f_1 + f_2 = 50$ que corresponde a este nodo.

Puede analizar el flujo a través de toda una red al construir dichas ecuaciones y resolver el sistema resultante de ecuaciones lineales.

Ejemplo 2.30

Describa los posibles flujos a través de la red de tuberías de agua que se muestra en la figura 2.11, donde el flujo se mide en litros por minuto.

Solución En cada nodo, escriba la ecuación que representa la conservación del flujo. Luego reescriba cada ecuación con las variables a la izquierda y la constante a la derecha, para obtener un sistema lineal en forma estándar.

$$\begin{array}{ll} \text{Node } A\text{:} & 15 = f_1 + f_4 \\ \text{Node } B\text{:} & f_1 = f_2 + 10 \\ \text{Node } C\text{:} & f_2 + f_3 + 5 = 30 \\ \text{Node } D\text{:} & f_4 + 20 = f_3 \end{array} \longrightarrow \begin{array}{rcrcrcrcl} f_1 & & & & & + & f_4 & = & 15 \\ f_1 & - & f_2 & & & & & = & 10 \\ & & f_2 & + & f_3 & & & = & 25 \\ & & & & f_3 & - & f_4 & = & 20 \end{array}$$

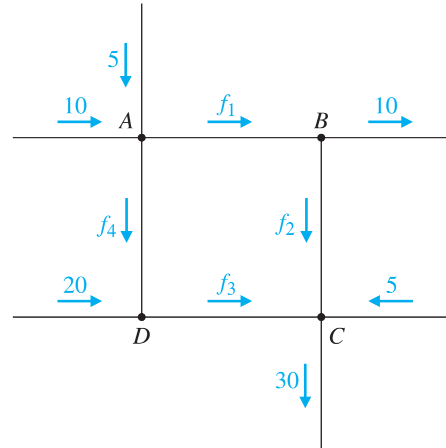


Figura 2.11

Con eliminación de Gauss-Jordan, reduzca la matriz aumentada:

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 15 \\ 1 & -1 & 0 & 0 & 10 \\ 0 & 1 & 1 & 0 & 25 \\ 0 & 0 & 1 & -1 & 20 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 15 \\ 0 & 1 & 0 & 1 & 5 \\ 0 & 0 & 1 & -1 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$



(Compruebe esto.) Se ve que existe una variable libre, f_4, de modo que se tiene un número infinito de soluciones. Al hacer $f_4 = t$ y expresar las variables pivote en términos de f_4, se obtiene

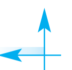
$$\begin{aligned} f_1 &= 15 - t \\ f_2 &= \ \ 5 - t \\ f_3 &= 20 + t \\ f_4 &= \qquad\ t \end{aligned}$$

Estas ecuaciones describen todos los flujos posibles y permiten analizar la red. Por ejemplo, se ve que, si se controla el flujo en la rama AD de modo que $t = 5$ L/min, entonces los otros flujos son $f_1 = 10$, $f_2 = 0$, y $f_3 = 25$.

Puede hacerlo todavía mejor. Puede encontrar los posibles flujos mínimo y máximo en cada rama. Cada uno de los flujos debe ser no negativo. Al examinar la primera y segunda ecuaciones a la vez, se ve que $t \leq 15$ (de otro modo f_1 sería negativo) y $t \leq 5$ (de otro modo f_2 sería negativo). La segunda de estas desigualdades es más restrictiva que la primera, de modo que debe usarla. La tercera ecuación no aporta más restricciones al parámetro t, así se deduce que $0 \leq t \leq 5$. Al combinar este resultado con las cuatro ecuaciones, se ve que

$$\begin{aligned} 10 &\leq f_1 \leq 15 \\ 0 &\leq f_2 \leq 5 \\ 20 &\leq f_3 \leq 25 \\ 0 &\leq f_4 \leq 5 \end{aligned}$$

Ahora se tiene una descripción completa de los posibles flujos a través de esta red.



Redes eléctricas

Las redes eléctricas son un tipo especializado de red que proporciona información acerca de fuentes de energía, como las baterías, y dispositivos impulsados por dichas fuentes, como las bombillas eléctricas o motores. Una fuente de energía "fuerza" a una corriente de electrones a fluir a través de la red, donde encuentran varios *resistores*, cada uno de los cuales requiere que se aplique cierta cantidad de fuerza para que la corriente fluya a través de él.

La ley fundamental de la electricidad es la ley de Ohm, que establece exactamente cuánta fuerza E se necesita para impulsar una corriente I a través de un resistor con resistencia R:

Ley de Ohm

$$\text{fuerza} = \text{resistencia} \times \text{corriente}$$

o

$$E = RI$$

La fuerza se mide en *volts*, la resistencia en *ohms* y la corriente en *amperes* (o *amps*, para abreviar). Por tanto, en términos de dichas unidades, la ley de Ohm se convierte en "volts = ohms × amps", y dice que es la "caída de voltaje" cuando una corriente pasa a través de un resistor; esto es: cuánto voltaje se usa.

La corriente fluye desde la terminal positiva de una batería y viaja de vuelta hacia la terminal negativa, recorriendo uno o más circuitos cerrados en el proceso. En un diagrama de red eléctrica, las baterías se representan como ⊣⊢ (donde la terminal positiva es la barra vertical más larga) y los resistores se representan como —\/\/\/—. Las siguientes dos leyes, cuyo descubrimiento se debe a Kirchhoff, gobiernan las redes eléctricas. La primera es una ley de "conservación de flujo" en cada nodo; la segunda es una ley de "balanceo de voltaje" alrededor de cada circuito.

Leyes de Kirchhoff

Ley de corriente (nodos)
La suma de las corrientes que fluyen hacia cualquier nodo es igual a la suma de las corrientes que salen de dicho nodo.

Ley de voltaje (circuitos)
La suma de las caídas de voltaje alrededor de cualquier circuito es igual al voltaje total alrededor del circuito (proporcionado por las baterías).

La figura 2.12 ilustra las leyes de Kirchhoff. En el inciso (a), la ley de corriente produce $I_1 = I_2 + I_3$ (o $I_1 - I_2 - I_3 = 0$, como se le escribirá); el inciso (b) produce $4I = 10$, donde se usó la ley de Ohm para calcular la caída de voltaje de $4I$ en el resistor. Al usar las leyes de Kirchhoff, puede establecer un sistema de ecuaciones lineales que permitirán determinar las corrientes en una red eléctrica.

Ejemplo 2.31

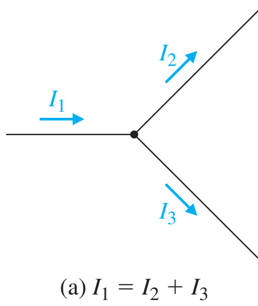
Determine las corrientes I_1, I_2 e I_3 en la red eléctrica que se muestra en la figura 2.13.

Solución Esta red tiene dos baterías y cuatro resistores. La corriente I_1 fluye a través de la rama superior *BCA*, la corriente I_2 fluye a través de la rama media *AB* y la corriente I_3 fluye a través de la rama inferior *BDA*.

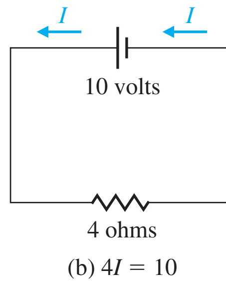
En el nodo *A*, la ley de corriente produce $I_1 + I_3 = I_2$ o

$$I_1 - I_2 + I_3 = 0$$

(Observe que se obtiene la misma ecuación en el nodo *B*.)



(a) $I_1 = I_2 + I_3$



(b) $4I = 10$

Figura 2.12

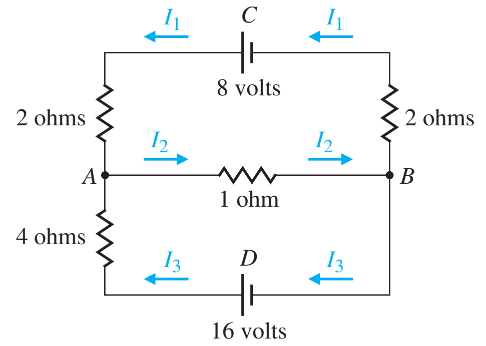


Figura 2.13

A continuación se aplica la ley de voltaje para cada circuito. Para el circuito *CABC*, las caídas de voltaje en los resistores son $2I_1$, I_2 y $2I_1$. Por tanto, se tiene la ecuación

$$4I_1 + I_2 = 8$$

De igual modo, para el circuito *DABD*, se obtiene

$$I_2 + 4I_3 = 16$$

(Note que en realidad hay un tercer circuito, *CADBC*, si va "contra la corriente". En este caso, debe tratar los voltajes y resistencias en las trayectorias "invertidas" como negativas. Al hacerlo se obtiene $2I_1 + 2I_1 - 4I_3 = 8 - 16 = -8$ o $4I_1 - 4I_3 = -8$, que se observa es justo la diferencia de las ecuaciones de voltaje para los otros dos circuitos. En consecuencia, puede omitir esta ecuación, pues no aporta nueva información. Por otra parte, incluirla no hace daño.)

Ahora se tiene un sistema de tres ecuaciones lineales con tres variables:

$$\begin{aligned} I_1 - I_2 + I_3 &= 0 \\ 4I_1 + I_2 \qquad &= 8 \\ I_2 + 4I_3 &= 16 \end{aligned}$$

La eliminación de Gauss-Jordan produce

$$\left[\begin{array}{rrr|r} 1 & -1 & 1 & 0 \\ 4 & 1 & 0 & 8 \\ 0 & 1 & 4 & 16 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 4 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

Por tanto, las corrientes son $I_1 = 1$ amp, $I_2 = 4$ amps e $I_3 = 3$ amps.

Comentario En algunas redes eléctricas, las corrientes pueden tener valores fraccionarios o incluso pueden ser negativas. Un valor negativo simplemente significa que la corriente en la rama correspondiente fluye en la dirección opuesta a la que se muestra en el diagrama de la red.

CAS

Ejemplo 2.32

La red que se muestra en la figura 2.14 tiene una sola fuente de poder *A* y cinco resistores. Encuentre las corrientes $I, I_1, \ldots, I_5$. Este es un ejemplo de lo que en ingeniería eléctrica se conoce como *circuito puente de Wheatstone*.

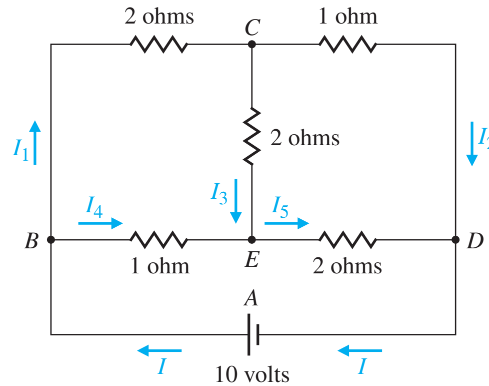


Figura 2.14
Un circuito puente

Solución La ley de corriente de Kirchhoff produce las siguientes ecuaciones en los cuatro nodos:

$$
\begin{aligned}
&\text{Node } B\text{:} \quad I - I_1 - I_4 = 0 \\
&\text{Node } C\text{:} \quad I_1 - I_2 - I_3 = 0 \\
&\text{Node } D\text{:} \quad I - I_2 - I_5 = 0 \\
&\text{Node } E\text{:} \quad I_3 + I_4 - I_5 = 0
\end{aligned}
$$

Para los tres circuitos básicos, la ley de voltaje produce

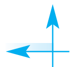
$$
\begin{aligned}
&\text{Circuit } ABEDA\text{:} \quad I_4 + 2I_5 = 10 \\
&\text{Circuit } BCEB\text{:} \quad 2I_1 + 2I_3 - I_4 = 0 \\
&\text{Circuit } CDEC\text{:} \quad I_2 - 2I_5 - 2I_3 = 0
\end{aligned}
$$

(Observe que la rama *DAB* no tiene resistor y por tanto no hay caída de voltaje; por ende, no hay término *I* en la ecuación para el circuito *ABEDA*. Note también que se cambiaron los signos tres veces, porque se avanzó "contra la corriente". Esto no plantea problemas, pues se dejará que el signo de la respuesta determine la dirección del flujo de corriente.)

Ahora se tiene un sistema de siete ecuaciones con seis variables. La reducción por renglones produce

$$
\left[\begin{array}{rrrrrr|r}
1 & -1 & 0 & 0 & -1 & 0 & 0 \\
0 & 1 & -1 & -1 & 0 & 0 & 0 \\
1 & 0 & -1 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 1 & 2 & 10 \\
0 & 2 & 0 & 2 & -1 & 0 & 0 \\
0 & 0 & 1 & -2 & 0 & -2 & 0
\end{array}\right]
\longrightarrow
\left[\begin{array}{rrrrrr|r}
1 & 0 & 0 & 0 & 0 & 0 & 7 \\
0 & 1 & 0 & 0 & 0 & 0 & 3 \\
0 & 0 & 1 & 0 & 0 & 0 & 4 \\
0 & 0 & 0 & 1 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 1 & 0 & 4 \\
0 & 0 & 0 & 0 & 0 & 1 & 3 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
$$

(Use su calculadora o CAS para comprobar esto.) Por tanto, la solución (en amps) es $I = 7, I_1 = I_5 = 3, I_2 = I_4 = 4$ e $I_3 = -1$. El significado del valor negativo aquí es que la corriente a través de la rama *CE* fluye en dirección opuesta a la que se marcó en el diagrama.



Comentario En este ejemplo sólo hay una fuente de poder, de modo que la única batería de 10 volts envía una corriente de 7 amps a través de la red. Si dichos valores se sustituyen en la ley de Ohm, $E = RI$, se obtiene $10 = 7R$ o $R = \frac{10}{7}$. Por tanto, toda la red

se comporta como si hubiese un solo resistor de $\frac{10}{7}$ ohms. Este valor se conoce como *resistencia efectiva* (R_{ef}) de la red.

Modelos económicos lineales

Una economía es un sistema muy complejo con muchas interrelaciones entre sus diversos sectores y los bienes y servicios que se producen y consumen. Determinar los precios óptimos y los niveles de producción sujetos a las metas económicas deseadas requiere sofisticados modelos matemáticos. El álgebra lineal resulta ser una poderosa herramienta en el desarrollo y análisis de tales modelos económicos.

En esta sección se introducen dos modelos basados en el trabajo del economista de Harvard, Wassily Leontief, en la décda de 1930. Sus métodos, a los que con frecuencia se les refiere como ***análisis input-output***, ahora son herramientas estándar en la economía matemática y los usan ciudades, corporaciones y países completos para planificación y pronósticos económicos.

Comience con un ejemplo simple.

Ejemplo 2.33

La economía de una región consiste de tres industrias o sectores: servicios, electricidad y producción de petróleo. Por simplicidad, se supone que cada industria produce un solo artículo (bienes o servicios) en un año dado y que el ***ingreso*** (***output***) se genera a partir de la venta de este artículo. Cada industria compra artículos de las otras industrias, incluida ella misma, para generar su output. Ningún artículo se compra fuera de la región y ninguna salida (output) se vende fuera de la región. Más aún, para cada industria se supone que la producción iguala exactamente al consumo (output igual a input, ingreso igual a gasto). En este sentido, esta es una ***economía cerrada*** que está en ***equilibrio***. La tabla 2.4 resume cuánta output de cada industria consume cada industria.

Tabla 2.4

		Producido por (output)		
		Servicios	**Electricidad**	**Petróleo**
Consumido por (input)	**Servicios**	1/4	1/3	1/2
	Electricidad	1/4	1/3	1/4
	Petróleo	1/2	1/3	1/4

Wassily Leontief (1906–1999) nació en San Petersburgo, Rusia. Estudió en la Universidad de Leningrado y recibió su doctorado de la Universidad de Berlín. Emigró a Estados Unidos en 1931 e impartió clases en la Universidad de Harvard y más tarde en la Universidad de New York. En 1932, Leontief comenzó a recopilar datos para la monumental tarea de realizar un análisis *input-output* de la economía de Estados Unidos, cuyos resultados se publicaron en 1941. Fue también uno de los primeros usuarios de computadoras, que necesitaba para resolver los sistemas lineales de gran escala de sus modelos. Por su trabajo pionero, Leontief recibió el Premio Nobel de Economía en 1973.

De la primera columna de la tabla se ve que la industria de servicios consume 1/4 de su propia output, la electricidad consume otro 1/4 y la industria petrolera usa 1/2 de la output de la industria de servicios. Las otras dos columnas tienen interpretaciones similares. Note que la suma de cada columna es 1, lo que indica que se consumen todas las output de cada industria.

Sean x_1, x_2 y x_3 que denotan el output (ingreso) anual de las industrias de servicios, electricidad y petróleo, respectivamente, en millones de dólares. Dado que el consumo corresponde al gasto, la industria de servicios gasta $\frac{1}{4}x_1$ en su propio artículo, $\frac{1}{3}x_2$ en electricidad y $\frac{1}{2}x_3$ en petróleo. Esto significa que el gasto total anual de la industria de servicios es $\frac{1}{4}x_1 + \frac{1}{3}x_2 + \frac{1}{2}x_3$. Dado que la economía está en equilibrio, el gasto de la industria

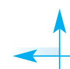
de servicios debe ser igual a su ingreso anual x_1. Esto genera la primera de las siguientes ecuaciones; las otras dos ecuaciones se obtienen al analizar los gastos de las industrias de electricidad y petróleo.

$$\begin{aligned} \text{Servicios:} \quad & \tfrac{1}{4}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{2}x_3 = x_1 \\ \text{Electricidad:} \quad & \tfrac{1}{4}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{4}x_3 = x_2 \\ \text{Petróleo:} \quad & \tfrac{1}{2}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{4}x_3 = x_3 \end{aligned}$$

Al reordenar cada ecuación, se obtiene un sistema homogéneo de ecuaciones lineales, que entonces se resuelven. (¡Compruebe esto!)

$$\begin{aligned} -\tfrac{3}{4}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{2}x_3 &= 0 \\ \tfrac{1}{4}x_1 - \tfrac{2}{3}x_2 - \tfrac{1}{4}x_3 &= 0 \\ \tfrac{1}{2}x_1 + \tfrac{1}{3}x_2 - \tfrac{3}{4}x_3 &= 0 \end{aligned} \longrightarrow \left[\begin{array}{ccc|c} -\frac{3}{4} & \frac{1}{3} & \frac{1}{2} & 0 \\ \frac{1}{4} & -\frac{2}{3} & \frac{1}{4} & 0 \\ \frac{1}{2} & \frac{1}{3} & -\frac{3}{4} & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -\frac{3}{4} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Al hacer $x_3 = t$, se encuentra que $x_1 = t$ y $x_2 = \frac{3}{4}t$. Por tanto, se ve que las outputs *relativas* de las industrias de servicios, electricidad y petrolera deben estar en las proporciones $x_1 : x_2 : x_3 = 4 : 3 : 4$ para que la economía esté en equilibrio.



Comentarios

- El último ejemplo ilustra lo que comúnmente se conoce como ***modelo cerrado de Leontief***.
- Dado que las output corresponden a ingreso, también puede considerar a x_1, x_2 y x_3 como los *precios* de los tres artículos.

Ahora se modifica el modelo del ejemplo 2.33 para acomodar una ***economía abierta***, una en la que haya tanto demanda *externa* como interna por los artículos que se producen. No es de sorprender que esta versión se conozca como ***modelo abierto de Leontief***.

Ejemplo 2.34

Considere las tres industrias del ejemplo 2.33, pero con consumo dado por la tabla 2.5. Se ve que, de los artículos producidos por la industria de servicios, 20% los consume la industria de servicios, 40% la industria eléctrica y 10% la industria petrolera. Por tanto, sólo 70% de la output de la industria de servicios lo consume esta economía. La implicación de este cálculo es que hay un exceso de output (ingreso) sobre el input (gasto) para la industria de servicios. Se dice que la industria de servicios es ***productiva***. Del mismo modo, la industria petrolera es productiva, pero la industria eléctrica es ***no productiva***. (Esto se refleja en el hecho de que las sumas de la primera y tercera columnas son menos que 1, pero la suma de la segunda columna es igual a 1.) El exceso de output puede aplicarse para satisfacer una demanda externa.

Tabla 2.5

		Producido por (output)		
		Servicios	**Electricidad**	**Petróleo**
Consumido por (input)	**Servicios**	0.20	0.50	0.10
	Electricidad	0.40	0.20	0.20
	Petróleo	0.10	0.30	0.30

Por ejemplo, suponga que hay una demanda externa anual (en millones de dólares) de 10, 10 y 30 de las industrias de servicios, eléctrica y petrolera, respectivamente. Entonces, al igualar los gastos (demanda interna y demanda externa) con el ingreso (output), se obtienen las siguientes ecuaciones:

	output	**demanda interna**	**demanda externa**
Servicios	x_1	$= 0.2x_1 + 0.5x_2 + 0.1x_3$	$+ 10$
Eléctrica	x_2	$= 0.4x_1 + 0.2x_2 + 0.2x_3$	$+ 10$
Petrolera	x_3	$= 0.1x_1 + 0.3x_2 + 0.3x_3$	$+ 30$

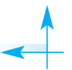
Al reordenar, se obtienen los siguientes sistemas lineal y matriz aumentada:

$$\begin{aligned} 0.8x_1 - 0.5x_2 - 0.1x_3 &= 10 \\ -0.4x_1 + 0.8x_2 - 0.2x_3 &= 10 \\ -0.1x_1 - 0.3x_2 + 0.7x_3 &= 30 \end{aligned} \longrightarrow \left[\begin{array}{rrr|r} 0.8 & -0.5 & -0.1 & 10 \\ -0.4 & 0.8 & -0.2 & 10 \\ -0.1 & -0.3 & 0.7 & 30 \end{array}\right]$$

CAS La reducción por renglón produce

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 61.74 \\ 0 & 1 & 0 & 63.04 \\ 0 & 0 & 1 & 78.70 \end{array}\right]$$

de donde se ve que las industrias de servicios, eléctrica y petrolera deben tener una producción anual de \$61.74, \$63.04 y \$78.70 (millones), respectivamente, para satisfacer las demandas interna y externa de sus artículos.



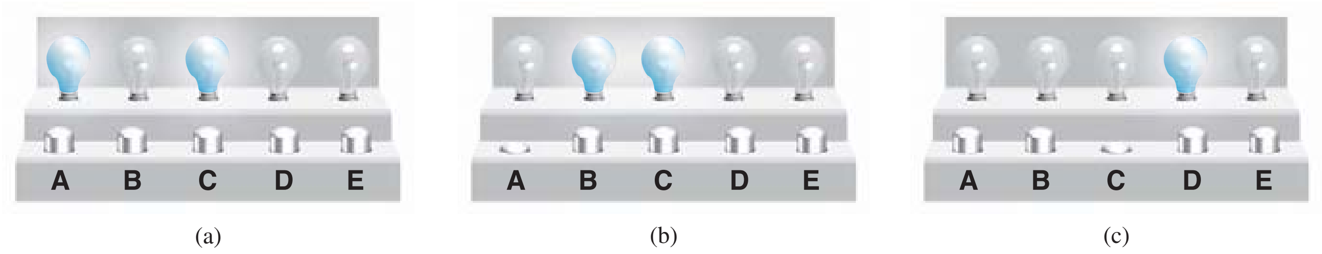
En la sección 3.7 se volverá a revisar estos modelos.

Juegos lineales finitos

Existen muchas situaciones en las que se debe considerar un sistema físico que sólo tenga un número finito de *estados*. En ocasiones, dichos estados pueden alterarse con la aplicación de ciertos procesos, cada uno de los cuales produce resultados finitos. Por ejemplo, una bombilla puede estar encendida o apagada, y un interruptor puede cambiar el estado de la bombilla de encendido a apagado y viceversa. Los sistemas digitales que surgen en ciencias de la computación con frecuencia son de este tipo. De manera más frívola, muchos juegos de computadora presentan acertijos en los que cierto dispositivo debe ser manipulado por varios interruptores para producir un resultado deseado. La finitud de tales situaciones es perfectamente adecuada para el análisis mediante aritmética modular, y con frecuencia los sistemas lineales sobre algún $\mathbb{Z}_p$ desempeñan un papel. Los problemas que involucran este tipo de situación con frecuencia se llaman ***juegos lineales finitos***.

Ejemplo 2.35

Una hilera de cinco luces se controla mediante cinco interruptores. Cada interruptor cambia el estado (encendido o apagado) de la luz directamente arriba de él y los estados de las luces inmediatamente adyacentes a izquierda y derecha. Por ejemplo, si la primera y tercera luces están encendidas, como en la figura 2.15(a), entonces oprimir el interruptor A cambia el estado del sistema al que se muestra en la figura 2.15(b). Si a continuación se oprime el interruptor C, entonces el resultado es el estado que se muestra en la figura 2.15(c).



(a) (b) (c)

Figura 2.15

Suponga que, inicialmente, todas las luces están apagadas. ¿Puede oprimir los interruptores en algún orden de modo que sólo la primera, tercera y quinta luces estén encendidas? ¿Puede oprimir los interruptores en algún orden de modo que sólo la primera luz esté encendida?

Solución La naturaleza encendido/apagado de este problema sugiere que la notación binaria será útil y que debe trabajar con $\mathbb{Z}_2$. En concordancia, los estados de las cinco luces se representan mediante un vector en $\mathbb{Z}_2^5$, donde 0 representa apagado y 1 representa encendido. Por tanto, por ejemplo, el vector

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

corresponde a la figura 2.15(b).

También puede usar vectores en $\mathbb{Z}_2^5$ para representar la acción de cada interruptor. Si un interruptor cambia el estado de una luz, el componente correspondiente es un 1; de otro modo, es 0. Con esta convención, las acciones de los cinco interruptores están dados por

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \ \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{d} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{e} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

La situación que se muestra en la figura 2.15(a) corresponde al estado inicial

$$\mathbf{s} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

seguido por

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Es la suma vectorial (en $\mathbb{Z}_2^5$)

$$\mathbf{s} + \mathbf{a} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

Observe que este resultado concuerda con la figura 2.15(b).

A partir de cualquier configuración inicial **s**, suponga que se oprimen los interruptores en el orden A, C, D, A, C, B. Esto corresponde a la suma vectorial $\mathbf{s} + \mathbf{a} + \mathbf{c} + \mathbf{d} + \mathbf{a} + \mathbf{c} + \mathbf{b}$. Pero, en $\mathbb{Z}_2^5$, la suma es conmutativa, de modo que se tiene

$$\begin{aligned} \mathbf{s} + \mathbf{a} + \mathbf{c} + \mathbf{d} + \mathbf{a} + \mathbf{c} + \mathbf{b} &= \mathbf{s} + 2\mathbf{a} + \mathbf{b} + 2\mathbf{c} + \mathbf{d} \\ &= \mathbf{s} + \mathbf{b} + \mathbf{d} \end{aligned}$$



donde se usó el hecho de que $2 = 0$ en $\mathbb{Z}_2$. Por ende, se lograría el mismo resultado al oprimir sólo B y D, y el orden no importa. (Compruebe que esto es correcto.) Por tanto, en este ejemplo no es necesario oprimir algún interruptor más de una vez.

De este modo, para ver si es posible lograr una configuración objetivo **t** a partir de una configuración inicial **s**, es necesario determinar si existen escalares $x_1, \ldots, x_5$ en $\mathbb{Z}_2$ tales que

$$\mathbf{s} + x_1\mathbf{a} + x_2\mathbf{b} + \cdots + x_5\mathbf{e} = \mathbf{t}$$

En otras palabras, es necesario resolver (si es posible) el sistema lineal sobre $\mathbb{Z}_2$ que corresponde a la ecuación vectorial

$$x_1\mathbf{a} + x_2\mathbf{b} + \cdots + x_5\mathbf{e} = \mathbf{t} - \mathbf{s} = \mathbf{t} + \mathbf{s}$$

En este caso, $\mathbf{s} = \mathbf{0}$ y la primera configuración objetivo es

$$\mathbf{t} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

La matriz aumentada de este sistema tiene como columnas los vectores dados:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array}\right]$$

Al reducir sobre $\mathbb{Z}_2$ se obtiene

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Por tanto, x_5 es una variable libre. En consecuencia, existen exactamente dos soluciones (que corresponden a $x_5 = 0$ y $x_5 = 1$). Al despejar las otras variables en términos de x_5, se obtiene

$$\begin{aligned} x_1 &= \phantom{1 + {}} x_5 \\ x_2 &= 1 + x_5 \\ x_3 &= 1 \\ x_4 &= 1 + x_5 \end{aligned}$$

De modo que, cuando $x_5 = 0$ y $x_5 = 1$, se tienen las soluciones

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$



respectivamente. (Compruebe que estas dos funcionan.)

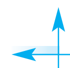
De igual modo, en el segundo caso se tiene

$$\mathbf{t} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

La matriz aumentada se reduce del modo siguiente:

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

lo que muestra que no hay solución en este caso; esto es: es imposible comenzar con todas las luces apagadas y sólo encender la primera luz.



El ejemplo 2.35 muestra el poder del álgebra lineal. Aun cuando haya encontrado mediante ensayo y error que no había solución, comprobar todas las formas posibles de oprimir los interruptores habría sido extremadamente tedioso. También podría haberse pasado por alto el hecho de que ningún interruptor necesita oprimirse más de una vez.

Ejemplo 2.36

Considere una hilera sólo con tres luces, cada una de las cuales puede estar apagado, ser azul claro o azul oscuro. Abajo de las luces hay tres interruptores, A, B y C, cada uno de los cuales cambia los estados de luces particulares al *siguiente* estado, en el orden que muestra la figura 2.16. El interruptor A cambia los estados de las primeras dos luces, el interruptor B las tres luces y el interruptor C las últimas dos luces. Si las tres luces inicialmente están apagadas, ¿es posible oprimir los interruptores en cierto orden de modo

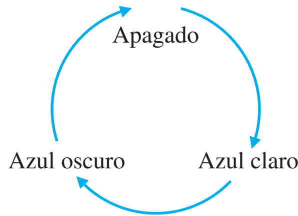


Figura 2.16

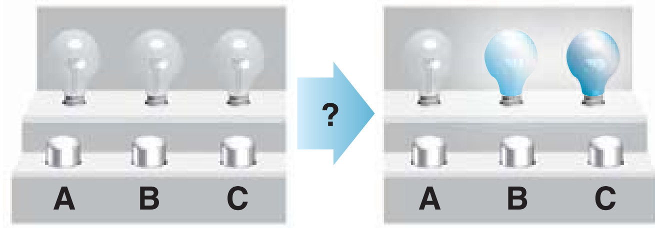


Figura 2.17

que las luces estén apagadas, azul claro y azul oscuro, en ese orden (como en la figura 2.17)?

Solución Mientras que el ejemplo 2.35 involucraba $\mathbb{Z}_2$, este claramente (¿es claro?) involucra $\mathbb{Z}_3$. En concordancia, los interruptores corresponden a los vectores

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \ \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \ \mathbf{c} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

en $\mathbb{Z}_3^3$, y la configuración final que se busca es $\mathbf{t} = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix}$. (Apagado es 0, azul claro es 1 y azul oscuro es 2.) Se quiere encontrar escalares x_1, x_2, x_3 en $\mathbb{Z}_3$ tales que

$$x_1\mathbf{a} + x_2\mathbf{b} + x_3\mathbf{c} = \mathbf{t}$$

(donde x_i representa el número de veces que se oprime el i-ésimo interruptor). Esta ecuación da lugar a la matriz aumentada $[\mathbf{a}\ \mathbf{b}\ \mathbf{c} \mid \mathbf{t}]$, que se reduce sobre $\mathbb{Z}_3$ del modo siguiente:

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array}\right]$$

Por tanto, existe una solución única: $x_1 = 2, x_2 = 1, x_3 = 1$. En otras palabras, debe oprimir el interruptor A dos veces y los otros dos interruptores una vez cada uno. (Compruébelo.)

Ejercicios 2.4

Asignación de recursos

1. Suponga que, en el ejemplo 2.27, en el tubo de ensayo se colocan cada día 400 unidades de alimento A, 600 unidades de B y 600 unidades de C, y que los datos del consumo diario de alimento por bacteria (en unidades por día) son los que se muestran en la tabla 2.6. ¿Cuántas bacterias de cada cepa pueden coexistir en el tubo de ensayo y consumir todo el alimento?

2. Suponga que, en el ejemplo 2.27, en el tubo de ensayo se colocan cada día 400 unidades de alimento A, 500 unidades de B y 600 unidades de C, y que los datos del consumo diario de alimento por bacteria (en unidades por día) son los que se muestran en la tabla 2.7. ¿Cuántas

Tabla 2.6

	Bacteria cepa I	Bacteria cepa II	Bacteria cepa III
Alimento A	1	2	0
Alimento B	2	1	1
Alimento C	1	1	2

Tabla 2.7

	Bacteria cepa I	Bacteria cepa II	Bacteria cepa III
Alimento A	1	2	0
Alimento B	2	1	3
Alimento C	1	1	1

bacterias de cada cepa pueden coexistir en el tubo de ensayo y consumir todo el alimento?

3. Una florista ofrece tres tamaños de arreglos florales que contienen rosas, margaritas y crisantemos. Cada arreglo pequeño contiene una rosa, tres margaritas y tres crisantemos. Cada arreglo mediano contiene dos rosas, cuatro margaritas y seis crisantemos. Cada arreglo grande contiene cuatro rosas, ocho margaritas y seis crisantemos. Un día, la florista nota que usó un total de 24 rosas, 50 margaritas y 48 crisantemos para surtir pedidos de estos tres tipos de arreglos. ¿Cuántos arreglos de cada tipo elaboró?

4. **(a)** En su bolsillo usted tiene algunas monedas de cinco, diez y 25 centavos. En total tiene 20 monedas y exactamente el doble de monedas de diez centavos que de cinco. El valor total de las monedas es $3.00. Encuentre el número de monedas de cada tipo.

(b) Encuentre *todas* las posibles combinaciones de 20 monedas (de cinco, diez y 25 centavos) que sumarían exactamente $3.00.

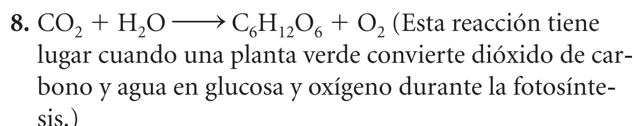
5. Un cafetalero vende tres mezclas de café. Una bolsa de la mezcla de la casa contiene 300 gramos de grano colombiano y 200 gramos de grano rostizado francés. Una bolsa de la mezcla especial contiene 200 gramos de grano colombiano, 200 gramos de grano keniano y 100 gramos de grano rostizado francés. Una bolsa de mezcla gourmet contiene 100 gramos de grano colombiano, 200 gramos de grano keniano y 200 gramos de grano rostizado francés. El comerciante tiene a la mano 30 kilogramos de grano colombiano, 15 kilogramos de grano keniano y 25 kilogramos de grano rostizado francés. Si quiere usar todos los granos, ¿cuántas bolsas de cada tipo de mezcla puede elaborar?

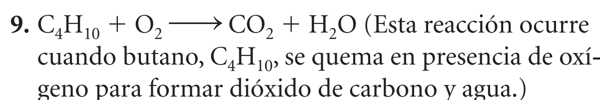
6. Vuelva a hacer el ejercicio 5 y suponga que la mezcla de la casa contiene 300 gramos de grano colombiano, 50 gramos de grano keniano y 150 gramos de grano rostizado francés, y que la mezcla gourmet contiene 100 gramos de grano colombiano, 350 gramos de grano keniano y 50 gramos de grano rostizado francés. Esta vez el comerciante tiene 30 kilogramos de grano colombiano, 15 kilogramos de grano keniano y 15 kilogramos de grano rostizado francés. Suponga que una bolsa de la mezcla de la casa produce una ganancia de $0.50, una bolsa de la mezcla especial produce una ganancia de $1.50 y una bolsa de la mezcla gourmet produce una ganancia de $2.00. ¿Cuántas bolsas de cada tipo debe preparar el comerciante si quiere usar todos los granos *y* maximizar su ganancia? ¿Cuál es la ganancia máxima?

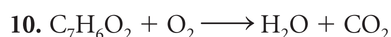
Balanceo de ecuaciones químicas

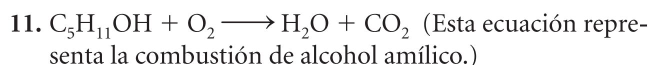
En los ejercicios 7-14, balancee la ecuación química para cada reacción.

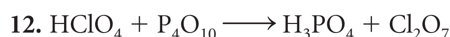

7. $FeS_2 + O_2 \longrightarrow Fe_2O_3 + SO_2$


8. $CO_2 + H_2O \longrightarrow C_6H_{12}O_6 + O_2$ (Esta reacción tiene lugar cuando una planta verde convierte dióxido de carbono y agua en glucosa y oxígeno durante la fotosíntesis.)

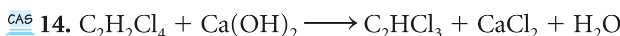

9. $C_4H_{10} + O_2 \longrightarrow CO_2 + H_2O$ (Esta reacción ocurre cuando butano, C_4H_{10}, se quema en presencia de oxígeno para formar dióxido de carbono y agua.)


10. $C_7H_6O_2 + O_2 \longrightarrow H_2O + CO_2$


11. $C_5H_{11}OH + O_2 \longrightarrow H_2O + CO_2$ (Esta ecuación representa la combustión de alcohol amílico.)


12. $HClO_4 + P_4O_{10} \longrightarrow H_3PO_4 + Cl_2O_7$


13. $Na_2CO_3 + C + N_2 \longrightarrow NaCN + CO$


CAS **14.** $C_2H_2Cl_4 + Ca(OH)_2 \longrightarrow C_2HCl_3 + CaCl_2 + H_2O$

Análisis de redes

15. La figura 2.18 muestra una red de tuberías con flujos medidos en litros por minuto.

(a) Establezca y resuelva un sistema de ecuaciones lineales para encontrar los flujos posibles.

(b) Si el flujo a través de *AB* se restringe a 5 L/min, ¿cuáles serán los flujos a través de las otras ramas?

(c) ¿Cuáles son los posibles flujos mínimo y máximo a través de cada rama?

(d) Se supuso que el flujo siempre es *positivo*. ¿Qué significaría flujo *negativo*, si supone que se permite? Proporcione una ilustración para este ejemplo.

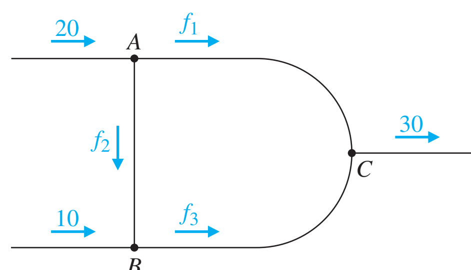


Figura 2.18

16. La parte central de Ciudad Gótica consiste de calles de un sentido, y el flujo de tráfico se midió en cada intersección. Para el bloque de ciudad que se muestra en la figura 2.19, los números representan la cantidad promedio de vehículos por minuto que entran y salen de las intersecciones A, B, C y D durante, las horas de oficina.

(a) Establezca y resuelva un sistema de ecuaciones lineales para encontrar los flujos posibles $f_1, \ldots, f_4$.

(b) Si el tráfico se regula en CD de modo que $f_4 = 10$ vehículos por minuto, ¿cuáles son los flujos promedio en las otras calles?

(c) ¿Cuáles son los posibles flujos mínimo y máximo en cada calle?

(d) ¿Cómo cambiaría la solución si *todas* las direcciones se invirtieran?

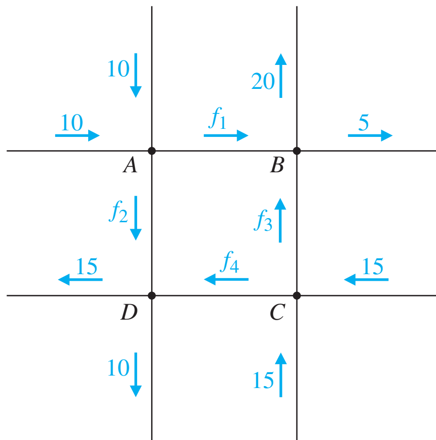


Figura 2.19

17. Una red de diques de irrigación se muestra en la figura 2.20, con flujos medidos en miles de litros por día.

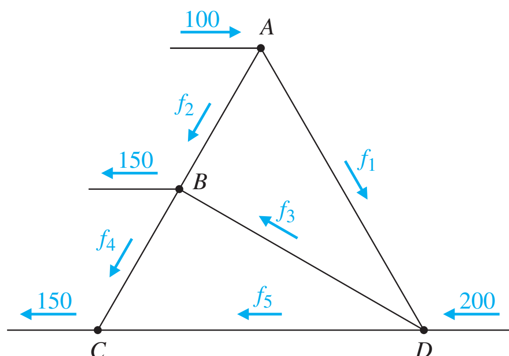


Figura 2.20

(a) Establezca y resuelva un sistema de ecuaciones lineales para encontrar los posibles flujos $f_1, \ldots, f_5$.

(b) Suponga que DC está cerrado. ¿Qué intervalo de flujo se necesitará mantener a través de DB?

(c) De la figura 2.20 es claro que DB no puede cerrarse. (¿Por qué no?) ¿Cómo su solución al inciso (a) demuestra esto?

(d) De su solución al inciso (a), determine los flujos mínimo y máximo a través de DB.

18. **(a)** Establezca y resuelva un sistema de ecuaciones lineales para encontrar los posibles flujos en la red que se muestra en la figura 2.21.

(b) ¿Es posible que $f_1 = 100$ y $f_6 = 150$? (Primero responda esta pregunta con referencia a su solución del inciso (a), y luego directamente de la figura 2.21.)

(c) Si $f_4 = 0$, ¿cuál será el rango de flujo en cada una de las otras ramas?

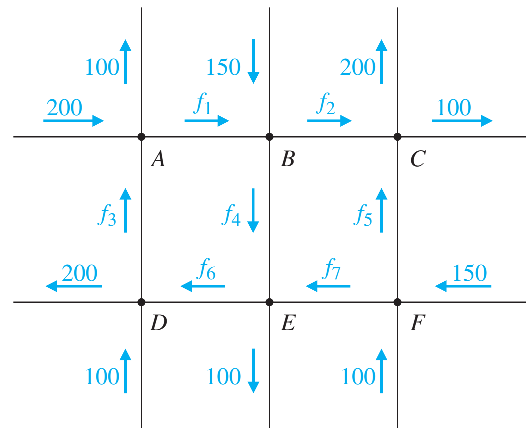
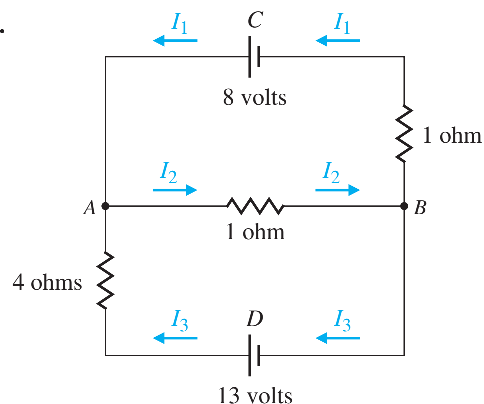


Figura 2.21

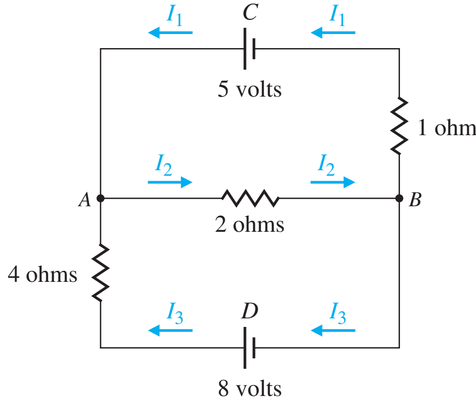
Redes eléctricas

Para los ejercicios 19 y 20, determine las corrientes para las redes eléctricas dadas.

19.



20.



21. **(a)** Encuentre las corrientes $I, I_1, \ldots, I_5$ en el circuito puente de la figura 2.22.

(b) Encuentre la resistencia efectiva de esta red.

(c) ¿Puede cambiar la resistencia en la rama *BC* (y dejar todo lo demás invariable) de modo que la corriente a través de la rama *CE* se vuelva 0?

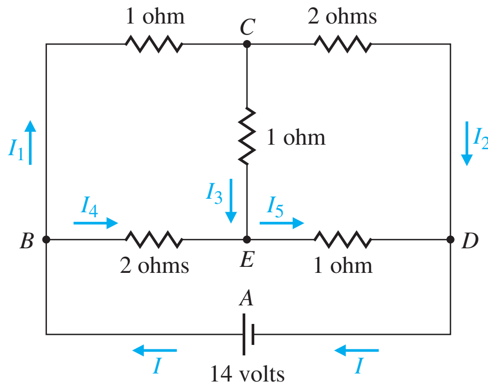


Figura 2.22

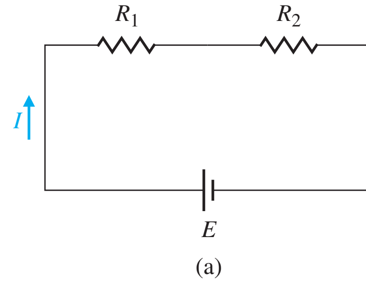
22. Las redes de los incisos (a) y (b) de la figura 2.23 muestran dos resistores acoplados en *serie* y en *paralelo*, respectivamente. Se quiere encontrar una fórmula general para la resistencia efectiva de cada red; esto es, encontrar R_{ef} de manera que $E = R_{ef}I$.

(a) Demuestre que la resistencia efectiva R_{ef} de una red con dos resistores acoplados en serie [figura 2.23(a)] está dada por

$$R_{ef} = R_1 + R_2$$

(b) Demuestre que la resistencia efectiva R_{ef} de una red con dos resistores acoplados en paralelo [figura 2.23(b)] está dada por

$$R_{eff} = \frac{1}{\dfrac{1}{R_1} + \dfrac{1}{R_2}}$$



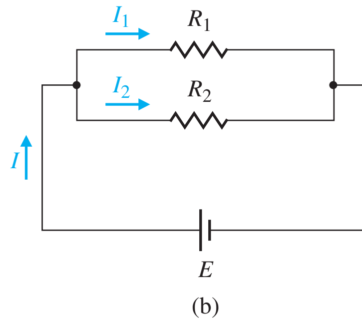


Figura 2.23
Resistores en serie y en paralelo

Modelos económicos lineales

23. Considere una economía simple con sólo dos industrias: agrícola y manufacturera. La agrícola consume 1/2 de los alimentos y 1/3 de los bienes manufacturados. La manufacturera consume 1/2 de los alimentos y 2/3 de los bienes manufacturados. Si supone que la economía está cerrada y en equilibrio, encuentre las salidas relativas (outputs) de las industrias agrícola y manufacturera.

24. Suponga que las industrias del carbón y el acero forman una economía cerrada. Cada \$1 producido por la industria del carbón requiere \$0.30 de carbón y \$0.70 de acero. Cada \$1 producido por la del acero requiere \$0.80 de carbón y \$0.20 de acero. Encuentre la producción anual (output) de carbón y acero si la producción anual total es de \$20 millones.

25. Un pintor, un plomero y un electricista entran a una cooperativa en la que cada uno de ellos está de acuerdo en trabajar para sí mismo y los otros dos por un total de 10 horas a la semana, de acuerdo con el horario que se muestra en la tabla 2.8. Para propósitos fiscales, cada persona debe establecer un valor para sus servicios. Están de acuerdo en hacerlo, de modo que cada uno de ellos quede a la par, es decir: de modo que el importe

total pagado por cada persona sea igual al importe que recibe. ¿Qué tarifa horaria debe cobrar cada persona si todas las tarifas son números enteros entre $30 y $60 por hora?

Tabla 2.8

		Proveedor		
		Pintor	**Plomero**	**Electricista**
	Pintor	2	1	5
Consumidor	**Plomero**	4	5	1
	Electricista	4	4	4

26. Cuatro vecinos, cada uno con un jardín de vegetales, está de acuerdo en compartir sus productos. Uno cosechará frijoles (B), otro lechuga (L), otro tomates (T) y otro calabazas (Z). La tabla 2.9 muestra qué fracción de cada cultivo recibirá cada vecino. ¿Qué precios deben cobrar los vecinos por sus cultivos, si cada persona debe quedar a la par y el cultivo de menor precio tiene un valor de $50?

Tabla 2.9

		Productor			
		B	**L**	**T**	**Z**
	B	0	1/4	1/8	1/6
Cosumidor	**L**	1/2	1/4	1/4	1/6
	T	1/4	1/4	1/2	1/3
	Z	1/4	1/4	1/8	1/3

27. Suponga que las industrias del carbón y del acero forman una economía abierta. Cada $1 producido por la industria del carbón requiere $0.15 de carbón y $0.20 de acero. Cada $1 producido por la del acero requiere $0.25 de carbón y $0.10 de acero. Suponga que hay una demanda exterior anual por $45 millones de carbón y $124 millones de acero.
(**a**) ¿Cuánto debe producir cada industria para satisfacer las demandas?
(**b**) Si la demanda de carbón disminuye en $5 millones al año, mientras que la demanda de acero aumenta en $6 millones al año, ¿cómo deben ajustar su producción las industrias del carbón y del acero?

28. En Ciudad Gótica, los departamentos de Administración (A), Salud (H) y Transporte (T) son independientes. Para cada dólar de servicios que producen, cada departamento usa cierta cantidad de servicios producidos por los otros departamentos y por él mismo, como se muestra en la tabla 2.10. Suponga que, durante el año, otros departamentos de la ciudad requieren $1 millón en servicios administrativos, $1.2 millones en servicios de salud y $0.8 millones en servicios de transporte. ¿Cuál debe ser el valor anual en dólares de los servicios producidos por cada departamento, con la finalidad de satisfacer las demandas?

Tabla 2.10

		Departmento		
		A	**H**	**T**
	A	$0.20	0.10	0.20
Compra	**H**	0.10	0.10	0.20
	T	0.20	0.40	0.30

Juegos lineales finitos

29. (**a**) En el ejemplo 2.35, suponga que todas las luces inicialmente están apagadas. ¿Puede oprimir los interruptores en algún orden para que sólo enciendan la segunda y cuarta luces?
(**b**) ¿Puede oprimir los interruptores en algún orden de modo que sólo se encienda la segunda luz?

30. (**a**) En el ejemplo 2.35, suponga que la cuarta luz inicialmente está encendida y que las otras cuatro luces están apagadas. ¿Puede oprimir los interruptores en algún orden de modo que sólo enciendan la segunda y cuarta luces?
(**b**) Puede oprimir las interruptores en algún orden de modo que sólo se encienda la segunda luz?

31. En el ejemplo 2.35, describa todas las posibles configuraciones de luces que pueden obtenerse si comienza con todas las luces apagadas.

32. (**a**) En el ejemplo 2.36, suponga que todas las luces inicialmente están apagadas. Demuestre que es posible presionar los interruptores en cierto orden de modo que las luces están apagadas, azul oscuro y azul claro, en ese orden.
(**b**) Demuestre que es posible oprimir los interruptores en cierto orden de modo que las luces estén azul claro, apagadas y azul oscuro, en ese orden.
(**c**) Pruebe que puede lograrse *cualquier* configuración de tres luces.

33. Suponga que las luces del ejemplo 2.35 pueden estar apagadas, azul claro o azul oscuro, y que los interruptores funcionan como se describe en el ejemplo 2.36. (Esto

es: los interruptores controlan las mismas luces que en el ejemplo 2.35, pero los colores avanzan en ciclo como en el ejemplo 2.36.) Demuestre que es posible comenzar con todas las luces apagadas y oprimir los interruptores en cierto orden de modo que las luces sean azul oscuro, azul claro, azul oscuro, azul claro y azul oscuro, en ese orden.

34. Para el ejercicio 33, describa todas las posibles configuraciones de luces que pueden obtenerse, iniciando con todas las luces apagadas.

CAS **35.** Nueve cuadrados, cada uno negro o blanco, se ordenan en una retícula de 3×3. La figura 2.24 muestra un posible arreglo.

1	2	3
4	5	6
7	8	9

Figura 2.24
El acertijo de nueve cuadrados

Cuando se toca, cada cuadrado cambia su estado y los estados de alguno de sus vecinos (negro $\rightarrow$ blanco y blanco $\rightarrow$ negro). La figura 2.25 muestra cómo funciona el cambio de estado. (Tocar el cuadrado cuyo número está en un círculo, hace que cambie el estado del cuadro marcado *.) El objeto del juego es volver negros los nueve cuadrados. [Los ejercicios 35 y 36 se adaptaron de acertijos que pueden encontrarse en el CD-ROM de juego interactivo *The Seventh Guest* (Trilobyte Software/Virgin Games, 1992).]

(a) Si la configuración inicial es la que se muestra en la figura 2.24, demuestre que puede ganar el juego y describa una secuencia de movimientos ganadora.

(b) Pruebe que el juego siempre puede ganarse, sin importar cuál sea la configuración inicial.

(1)*	2*	3
4*	5*	6
7	8	9

1*	(2)*	3*
4	5	6
7	8	9

1	2*	(3)*
4	5*	6*
7	8	9

1*	2	3
(4)*	5	6
7*	8	9

1	2*	3
4*	(5)*	6*
7	8*	9

1	2	3*
4	5	(6)*
7	8	9*

1	2	3
4*	5*	6
(7)*	8*	9

1	2	3
4	5	6
7*	(8)*	9*

1	2	3
4	5*	6*
7	8*	(9)*

Figura 2.25
Cambios de estado para el acertijo de nueve cuadrados

CAS **36.** Considere una variación del acertijo de nueve cuadrados. El juego es el mismo que se describe en el ejercicio 35, excepto que hay tres posibles estados para cada cuadrado: blanco, gris o negro. Los cuadrados cambian como se muestra en la figura 2.25, pero ahora el estado cambia siguiendo el ciclo blanco $\rightarrow$ gris $\rightarrow$ negro $\rightarrow$ blanco. Demuestre cómo puede lograrse la combinación ganadora todos negros a partir de la configuración inicial que se muestra en la figura 2.26.

1	2	3
4	5	6
7	8	9

Figura 2.26
Acertijo de nueve cuadrados con más estados

Problemas varios

En los ejercicios 37-53, establezca y resuelva un sistema adecuado de ecuaciones lineales para responder las preguntas.

37. Grace es tres veces mayor que Hans, pero en 5 años tendrá el doble de la edad que Hans tiene ahora. ¿Qué edades tienen ahora?

38. La suma de las edades de Annie, Bert y Chris es 60. Annie es mayor que Bert por el mismo número de años que Bert es mayor que Chris. Cuando Bert sea tan viejo como ahora es Annie, Annie será tres veces más vieja de lo que Chris es ahora. ¿Cuáles son sus edades?

Los dos problemas anteriores son típicos de los que se encuentran en libros populares de acertijos matemáticos. Sin embargo, tienen sus orígenes en la antigüedad. Una tablilla de arcilla de Babilonia, que sobrevive desde aproximadamente 300 a.C., contiene el siguiente problema.

39. Existen dos campos cuya área total es 1800 yardas cuadradas. Un campo produce grano a razón de 2/3 de fanega por yarda cuadrada; el otro campo produce grano a razón de $\frac{1}{2}$ fanega por yarda cuadrada. Si la producción total es de 1100 fanegas, ¿cuál es el tamaño de cada campo?

Hace más de 2000 años, los chinos desarrollaron métodos para resolver sistemas de ecuaciones lineales, incluida una versión de la eliminación gaussiana que no fue muy conocida en Europa sino hasta el siglo XIX. (No hay evidencia de que Gauss estuviese al tanto de los métodos chinos cuando desarrolló lo que ahora se conoce como eliminación gaussiana. Sin embargo, es claro que los chinos conocían la esencia del método, aun cuando no justificaron su uso.) El siguiente problema se toma del texto chino Jiuzhang suanshu (Nueve capítulos en el arte matemático), *escrito durante la temprana dinastía Han, alrededor de 200 a.C.*

40. Existen tres tipos de maíz. Tres gavillas del primer tipo, dos del segundo y una del tercero constituyen 39 medidas. Dos gavillas del primer tipo, tres del segundo y una del tercero constituyen 34 medidas. Y una gavilla del primer tipo, dos del segundo y tres del tercero constituyen 26 medidas. ¿Cuántas medidas de maíz contiene una gavilla de cada tipo?

41. Describa todos los posibles valores de a, b, c y d que harán que cada una de las siguientes sea una tabla de sumas válida. [Los problemas 41-44 se basan en el artículo “An Application of Matrix Theory” de Paul Glaister en *The Mathematics Teacher*, 85 (1992), pp. 220-223.]

(a)

+	a	b
c	2	3
d	4	5

(b)

+	a	b
c	3	6
d	4	5

42. ¿Qué condiciones de w, x, y y z garantizarán que puedan encontrarse a, b, c y d de modo que la siguiente sea una tabla de sumas válida?

+	a	b
c	w	x
d	y	z

43. Describa todos los posibles valores de a, b, c, d, e y f que harán de cada una de las siguientes una tabla de sumas válida.

(a)

+	a	b	c
d	3	2	1
e	5	4	3
f	4	3	1

(b)

+	a	b	c
d	1	2	3
e	3	4	5
f	4	5	6

44. Al generalizar el ejercicio 42, encuentre las condiciones para las entradas de una tabla de sumas 3×3 que garantizarán que pueda resolverse para a, b, c, d, e y f como antes.

45. De la geometría elemental se sabe que hay una sola línea recta que pasa a través de dos puntos cualesquiera en un plano. Menos conocido es el hecho de que existe una parábola única a través de cualesquiera *tres* puntos no colineales en un plano. Para cada de los siguientes conjuntos de puntos, encuentre una parábola con una ecuación de la forma $y = ax^2 + bx + c$ que pase a través de los puntos dados. (Grafique la parábola resultante para comprobar la validez de su respuesta.)

(a) $(0, 1), (-1, 4)$ y $(2, 1)$
(b) $(-3, 1), (-2, 2)$ y $(-1, 5)$

46. A través de cualesquiera tres puntos no colineales también pasa una circunferencia única. Encuentre las circunferencias (cuyas ecuaciones generales sean de la forma $x^2 + y^2 + ax + by + c = 0$) que pasan a través de los conjuntos de puntos del ejercicio 45. (Para comprobar la validez de su respuesta, encuentre el centro y el radio de cada circunferencia y dibuje una gráfica.)

El proceso de sumar funciones racionales (razones de polinomios) al colocarlos sobre un denominador común es el análogo de sumar números racionales. El proceso inverso, de descomponer una función racional al escribirla como una suma de funciones racionales más simples, es útil en muchas áreas de las matemáticas; por ejemplo, surge en cálculo cuando se debe integrar una función racional y en matemáticas discretas cuando se usan funciones generadoras para resolver relaciones de recurrencia. La descomposición de una función racional como una suma de fracciones parciales conduce a un sistema de ecuaciones lineales. En los ejercicios 47-50, encuentre la descomposición en fracciones parciales de la forma dada. (Las letras mayúsculas denotan constantes.)

47. $\dfrac{3x + 1}{x^2 + 2x - 3} = \dfrac{A}{x - 1} + \dfrac{B}{x + 3}$

48. $\dfrac{x^2 - 3x + 3}{x^3 + 2x^2 + x} = \dfrac{A}{x} + \dfrac{B}{x + 1} + \dfrac{C}{(x + 1)^2}$

CAS **49.** $$\frac{x - 1}{(x + 1)(x^2 + 1)(x^2 + 4)} = \frac{A}{x + 1} + \frac{Bx + C}{x^2 + 1} + \frac{Dx + E}{x^2 + 4}$$

CAS **50.** $$\frac{x^3 + x + 1}{x(x - 1)(x^2 + x + 1)(x^2 + 1)^3} = \frac{A}{x} + \frac{B}{x - 1} + \frac{Cx + D}{x^2 + x + 1} + \frac{Ex + F}{x^2 + 1} + \frac{Gx + H}{(x^2 + 1)^2} + \frac{Ix + J}{(x^2 + 1)^3}$$

A continuación hay dos fórmulas útiles para las sumas de potencias de números naturales consecutivos:

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

y

$$1^2 + 2^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

La validez de estas fórmulas para todos los valores de $n \geq 1$ *(o incluso* $n \geq 0$*) puede establecerse usando inducción matemática (vea el Apéndice B). Sin embargo, una forma de hacer una suposición educada acerca de cuáles son las fórmulas, es observar que puede reescribir las dos fórmulas anteriores como*

$$\tfrac{1}{2}n^2 + \tfrac{1}{2}n \quad \text{y} \quad \tfrac{1}{3}n^3 + \tfrac{1}{2}n^2 + \tfrac{1}{6}n$$

respectivamente. Esto conduce a la conjetura de que la suma de las p-ésimas potencias de los primeros números naturales es un polinomio de grado $p + 1$ *en la variable* n*.*

51. Si supone que $1 + 2 + \cdots + n = an^2 + bn + c$, encuentre a, b y c al sustituir tres valores para n y por tanto obtener un sistema de ecuaciones lineales con a, b y c.

52. Suponga que $1^2 + 2^2 + \cdots + n^2 = an^3 + bn^2 + cn + d$. Encuentre a, b, c y d. [*Sugerencia:* es legítimo usar $n = 0$. ¿Cuál es el lado izquierdo en este caso?]

53. Demuestre que $1^3 + 2^3 + \cdots + n^3 = (n(n+1)/2)^2$.

El sistema de posicionamiento global

El sistema de posicionamiento global (GPS, por sus siglas en inglés) se usa en muchas situaciones para determinar ubicaciones geográficas. Militares, topógrafos, líneas aéreas, compañías de transporte y excursionistas lo usan. La tecnología GPS se ha convertido en lugar común, a tal grado que algunos automóviles, teléfonos celulares y varios dispositivos manuales ahora están equipados con él.

La idea básica del GPS es una variante de la triangulación tridimensional: un punto sobre la superficie de la Tierra está determinado de manera exclusiva al conocer su distancia desde otros tres puntos. Aquí el punto que se quiere determinar es la ubicación del receptor GPS, los otros puntos son satélites y las distancias se calculan mediante los tiempos de recorrido de señales de radio enviadas de los satélites al receptor.

Se supondrá que la Tierra es una esfera sobre la cual se superpone un sistema de coordenadas *xyz*, con el centro de la Tierra en el origen y con el eje *z* positivo que corre a lo largo del polo norte y está fijo en relación con la Tierra.

Por simplicidad, considere que una unidad es igual al radio de la Tierra. Por tanto, la superficie de la Tierra se convierte en la esfera unitaria con ecuación $x^2 + y^2 + z^2 = 1$. El tiempo se medirá en centésimas de segundo. El GPS encuentra las distancias al conocer cuánto tarda una señal de radio en ir desde un punto a otro. Para esto es necesario conocer la rapidez de la luz, que es aproximadamente igual a 0.47 (radios terrestres por centésimas de segundo).

Imagine que usted es un excursionista perdido en el bosque en el punto (x, y, z) en cierto momento t. No sabe dónde está y, más aún, no tiene reloj, así que no sabe qué hora es. Sin embargo, tiene su dispositivo GPS y recibe señales simultáneas de cuatro satélites, que le dan sus posiciones y tiempos como se muestra en la tabla 2.11. (Las distancias se miden en radios terrestres y el tiempo en centésimas de segundo después de medianoche.)

Esta aplicación se basa en el artículo "An Underdetermined Linear System for GPS" por Dan Kalman en *The College Mathematics Journal*, 33 (2002), pp. 384-290. Para un tratamiento más profundo de las ideas aquí presentadas, vea G. Strang y K. Borre, *Linear Algebra, Geodesy, and GPS* (Wellesley-Cambridge Press, MA, 1997).

Tabla 2.11 Datos satelitales

Satélite	Posición	Tiempo
1	(1.11, 2.55, 2.14)	1.29
2	(2.87, 0.00, 1.43)	1.31
3	(0.00, 1.08, 2.29)	2.75
4	(1.54, 1.01, 1.23)	4.06



Sea (x, y, z) su posición, y sea t el tiempo cuando llegan las señales. La meta es resolver para x, y, z y t. Su distancia al satélite 1 puede calcularse del modo siguiente. La señal, que viaja a una rapidez de 0.47 radios terrestres/10^{-2} s, se envió en el tiempo 1.29 y llegó en el tiempo t, de modo que tardó $t - 1.29$ centésimas de segundo en llegar a usted. La distancia es igual a la velocidad multiplicada por el tiempo (transcurrido), de modo que

$$d = 0.47(t - 1.29)$$

También puede expresar d en términos de (x, y, z) y la posición del satélite (1.11, 2.55, 2.14) con el uso de la fórmula de distancia:

$$d = \sqrt{(x - 1.11)^2 + (y - 2.55)^2 + (z - 2.14)^2}$$

Combinar estos resultados conduce a la ecuación

$$(x - 1.11)^2 + (y - 2.55)^2 + (z - 2.14)^2 = 0.47^2(t - 1.29)^2 \qquad (1)$$

Al desarrollar, simplificar y reordenar, se encuentra que la ecuación (1) se convierte en

$$2.22x + 5.10y + 4.28z - 0.57t = x^2 + y^2 + z^2 - 0.22t^2 + 11.95$$

De igual modo, puede obtener una ecuación correspondiente para cada uno de los otros tres satélites. Termina con un sistema de cuatro ecuaciones con x, y, z y t:

$$\begin{aligned}
2.22x + 5.10y + 4.28z - 0.57t &= x^2 + y^2 + z^2 - 0.22t^2 + 11.95 \\
5.74x \qquad\qquad + 2.86z - 0.58t &= x^2 + y^2 + z^2 - 0.22t^2 + 9.90 \\
2.16y + 4.58z - 1.21t &= x^2 + y^2 + z^2 - 0.22t^2 + 4.74 \\
3.08x + 2.02y + 2.46z - 1.79t &= x^2 + y^2 + z^2 - 0.22t^2 + 1.26
\end{aligned}$$

No son ecuaciones lineales, pero los términos no lineales son los mismos en cada ecuación. Si resta la primera ecuación de cada una de las otras tres ecuaciones, obtiene un sistema lineal:

$$\begin{aligned}
3.52x - 5.10y - 1.42z - 0.01t &= 2.05 \\
-2.22x - 2.94y + 0.30z - 0.64t &= 7.21 \\
0.86x - 3.08y - 1.82z - 1.22t &= -10.69
\end{aligned}$$

La matriz aumentada se reduce por renglones a

$$\left[\begin{array}{cccc|c} 3.52 & -5.10 & -1.42 & -0.01 & -2.05 \\ -2.22 & -2.94 & 0.30 & -0.64 & -7.21 \\ 0.86 & -3.08 & -1.82 & -1.22 & -10.69 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0.36 & 2.97 \\ 0 & 1 & 0 & 0.03 & 0.81 \\ 0 & 0 & 1 & 0.79 & 5.91 \end{array}\right]$$

De donde se ve que

$$
\begin{aligned}
x &= 2.97 - 0.36t \\
y &= 0.81 - 0.03t \qquad (2) \\
z &= 5.91 - 0.79t
\end{aligned}
$$

con t libre. Al sustituir estas ecuaciones en (1) se obtiene

$$
\begin{aligned}
&(2.97 - 0.36t - 1.11)^2 + (0.81 - 0.03t - 2.55)^2 \\
&\quad + (5.91 - 0.79t - 2.14)^2 = 0.47^2(t - 1.29)^2
\end{aligned}
$$

que se simplifica en la ecuación cuadrática

$$0.54t^2 - 6.65t + 20.32 = 0$$

Existen dos soluciones:

$$t = 6.74 \quad \text{y} \quad t = 5.60$$

Al sustituir en (2), se encuentra que la primera solución corresponde a $(x, y, z) = (0.55, 0.61, 0.56)$ y la segunda solución a $(x, y, z) = (0.96, 0.65, 1.46)$. Claramente, la segunda solución no está sobre la esfera unitaria (Tierra), de modo que se le rechaza. La primera solución produce $x^2 + y^2 + z^2 = 0.99$, así que se satisface que, dentro de un error de redondeo aceptable, sus coordenadas se ubican como (0.55, 0.61, 0.56).

En la práctica, los GPS toman en cuenta muchos más factores, como el hecho de que la superficie de la Tierra no es exactamente esférica, de modo que se necesitan refinamientos adicionales que involucran técnicas como la aproximación por mínimos cuadrados (vea el capítulo 7). Además, los resultados de los cálculos GPS se convierten de coordenadas rectangulares (cartesianas) a latitud y longitud, un ejercicio interesante en sí mismo y que involucra incluso a otras ramas de las matemáticas.

CAS

2.5 Métodos iterativos para resolver sistemas lineales

Los métodos directos para resolver sistemas lineales, con el uso de operaciones elementales de renglón, conducen a soluciones exactas en muchos casos, pero están sujetas a errores debido a factores de redondeo y otros, como se ha visto. La tercera vía en el "trivium" lleva por una ruta de hecho muy diferente. En esta sección se exploran los métodos que proceden *iterativamente* al generar sucesivamente sucesiones de vectores que aproximan una solución para un sistema lineal. En muchas instancias (como cuando la matriz coeficiente es *dispersa*, esto es: contiene muchas entradas cero), los métodos iterativos pueden ser más rápidos y más precisos que los métodos directos. Además, los métodos iterativos pueden detenerse siempre que la solución aproximada que generen sea suficientemente precisa. Asimismo, los métodos iterativos con frecuencia se *benefician* de la imprecisión: en realidad, el error de redondeo puede acelerar su convergencia hacia una solución.

Se explorarán dos métodos iterativos para resolver sistemas lineales: el ***método de Jacobi*** y un refinamiento del mismo, el ***método de Gauss-Seidel***. En todos los ejemplos se considerarán sistemas lineales con el mismo número de variables que de ecuaciones y se supondrá que hay una solución única. El interés es encontrar esta solución mediante métodos iterativos.

Ejemplo 2.37

Considere el sistema

$$\begin{aligned} 7x_1 - x_2 &= 5 \\ 3x_1 - 5x_2 &= -7 \end{aligned}$$

El método de Jacobi comienza con la resolución de la primera ecuación para x_1 y la segunda ecuación para x_2, para obtener

$$\begin{aligned} x_1 &= \frac{5 + x_2}{7} \\ x_2 &= \frac{7 + 3x_1}{5} \end{aligned} \tag{1}$$

Ahora se necesita una ***aproximación inicial*** a la solución. Es evidente que no importa cuál sea esta aproximación inicial, así que podría tomarse $x_1 = 0, x_2 = 0$. Estos valores se usan en las ecuaciones (1) para obtener nuevos valores de x_1 y x_2:

$$\begin{aligned} x_1 &= \frac{5 + 0}{7} = \frac{5}{7} \approx 0.714 \\ x_2 &= \frac{7 + 3 \cdot 0}{5} = \frac{7}{5} = 1.400 \end{aligned}$$

Ahora se sustituyen estos valores en (1) para obtener

$$\begin{aligned} x_1 &= \frac{5 + 1.4}{7} \approx 0.914 \\ x_2 &= \frac{7 + 3 \cdot \frac{5}{7}}{5} \approx 1.829 \end{aligned}$$

(escritas con tres lugares decimales). Este proceso se repite (con los anteriores valores de x_2 y x_1 para obtener nuevos valores de x_1 y x_2), lo que produce la sucesión de aproximaciones dadas en la tabla 2.12.

Carl Gustav Jacobi (1804-1851) fue un matemático alemán que realizó importantes aportaciones a muchos campos de las matemáticas y la física, incluidas geometría, teoría de números, análisis, mecánica y dinámica de fluidos. Aunque gran parte de su trabajo fue en matemáticas aplicadas, Jacobi creía en la importancia de hacer matemáticas puras. Excelente profesor, dio cátedra en las Universidades de Berlín y Königsberg, y fue uno de los más famosos matemáticos de Europa.

Tabla 2.12

n	0	1	2	3	4	5	6
x_1	0	0.714	0.914	0.976	0.993	0.998	0.999
x_2	0	1.400	1.829	1.949	1.985	1.996	1.999

Los vectores sucesivos $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ se llaman ***iterados***, de modo que, por ejemplo, cuando $n = 4$, el cuarto iterado es $\begin{bmatrix} 0.993 \\ 1.985 \end{bmatrix}$. Puede ver que los iterados de este ejemplo se aproximan a $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, que es la solución exacta del sistema dado. (Compruébelo.)

En este caso se dice que el método de Jacobi ***converge***.

El método de Jacobi calcula los iterados sucesivos en un sistema de dos variables, de acuerdo con el patrón cruzado que se muestra en la tabla 2.13.

Tabla 2.13

n	0	1	2	3
x_1				
x_2				

El método de Gauss-Seidel recibe su nombre en honor de C. F. Gauss y Philipp Ludwig von Seidel (1821–1896). Seidel trabajó en análisis, teoría de probabilidad, astronomía y óptica. Por desgracia, sufrió de problemas oculares y se retiró a una edad joven. El ensayo en el que describió el método ahora conocido como de Gauss-Seidel se publicó en 1874. Gauss, al parecer, ¡no estaba al tanto del método!

Antes de considerar el método de Jacobi en el caso general, observe una modificación del mismo que con frecuencia converge con más rapidez a la solución. El *método de Gauss-Seidel* es igual que el método de Jacobi, excepto que usa cada nuevo valor *tan pronto como se puede*. De este modo, en el ejemplo, comience por calcular $x_1 = (5 + 0)/7 = \frac{5}{7} \approx 0.714$ como antes, pero ahora use este valor de x_1 para obtener el siguiente valor de x_2:

$$x_2 = \frac{7 + 3 \cdot \frac{5}{7}}{5} \approx 1.829$$

Luego use este valor de x_2 para recalcular x_1 y así sucesivamente. Esta vez los iterados se muestran en la tabla 2.14.

Observe que el método de Gauss-Seidel converge más rápido hacia la solución. Esta vez los iterados se calculan de acuerdo con el patrón en zigzag que se muestra en la tabla 2.15.

Tabla 2.14

n	0	1	2	3	4	5
x_1	0	0.714	0.976	0.998	1.000	1.000
x_2	0	1.829	1.985	1.999	2.000	2.000

Tabla 2.15

n	0	1	2	3
x_1				
x_2				

El método de Gauss-Seidel también tiene una agradable interpretación geométrica en el caso de dos variables. Puede considerar a x_1 y x_2 como las coordenadas de puntos en el plano. El punto de partida es el punto correspondiente a la aproximación inicial (0, 0). El primer cálculo produce $x_1 = \frac{5}{7}$, de modo que se avanza al punto $(\frac{5}{7}, 0) \approx (0.714, 0)$. Después se calcula $x_2 = \frac{64}{35} \approx 1.829$, que avanza hacia el punto $(\frac{5}{7}, \frac{64}{35}) \approx (0.714, 1.829)$. Al continuar de esta forma, los cálculos del método de Gauss-Seidel dan lugar a una sucesión de puntos, cada uno de los cuales difiere del punto anterior en exactamente una coordenada. Si se trazan las rectas $7x_1 - x_2 = 5$ y $3x_1 - 5x_2 = -7$ correspondientes a las dos ecuaciones dadas, se descubre que los puntos calculados arriba caen alternativamente en las dos rectas, como se muestra en la figura 2.27. Más aún, se aproximan al punto de intersección de las rectas, que corresponde a la solución del sistema de ecuaciones. ¡Esto es lo que significa *convergencia*!

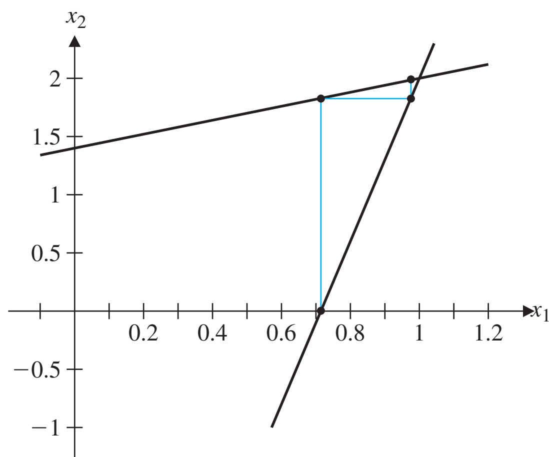


Figura 2.27
Iterados convergentes

Los casos generales de los dos métodos son análogos. Dado un sistema de n ecuaciones lineales con n variables,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned} \tag{2}$$

la primera ecuación se resuelve para x_1, la segunda para x_2, etcétera. Entonces, al comenzar con una aproximación inicial, se usan estas nuevas ecuaciones para actualizar

iterativamente cada variable. El método de Jacobi usa *todos* los valores en la k-ésima iteración para calcular el $(k + 1)$-ésimo iterado, mientras que el método de Gauss-Seidel siempre usa el valor *más reciente* de cada variable en cada cálculo. El ejemplo 2.39 más adelante ilustra el método de Gauss-Seidel en un problema de tres variables.

En este punto, acaso tenga algunas preguntas y preocupaciones acerca de estos métodos iterativos. (¿Las tiene?) Muchas vienen a la mente: ¿estos métodos deben converger? Si no, ¿*cuándo* convergen? *Si* convergen, ¿deben converger a la solución? La respuesta a la primera pregunta es no, como ilustra el ejemplo 2.38.

Ejemplo 2.38

Aplique el método de Gauss-Seidel al sistema

$$\begin{aligned} x_1 - x_2 &= 1 \\ 2x_1 + x_2 &= 5 \end{aligned}$$

con aproximación inicial $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

Solución Al reordenar las ecuaciones se obtiene

$$\begin{aligned} x_1 &= 1 + x_2 \\ x_2 &= 5 - 2x_1 \end{aligned}$$



Los primeros iterados se proporcionan en la tabla 2.16. (Compruébelos.)

La solución real al sistema dado es $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Claramente, los iterados en la tabla 2.16 no se aproximan a este punto, como la figura 2.28 deja gráficamente claro en un ejemplo de ***divergencia***.

Tabla 2.16

n	0	1	2	3	4	5
x_1	0	1	4	−2	10	−14
x_2	0	3	−3	9	−15	33

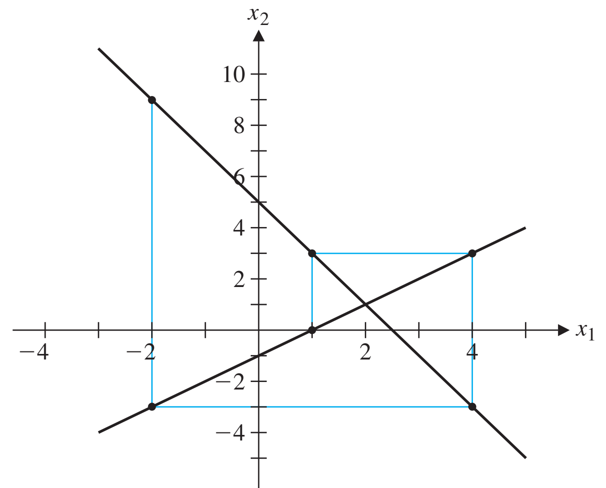
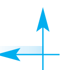


Figura 2.28
Iterados divergentes



De modo que, ¿cuándo convergen estos métodos iterativos? Por desgracia, la respuesta a esta pregunta es más bien complicada. Se le responderá por completo en el capítulo 7, pero por ahora se dará una respuesta parcial, sin demostración.

Sea A la matriz $n \times n$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Se dice que A es ***estrictamente diagonal dominante*** si

$$\begin{aligned} |a_{11}| &> |a_{12}| + |a_{13}| + \cdots + |a_{1n}| \\ |a_{22}| &> |a_{21}| + |a_{23}| + \cdots + |a_{2n}| \\ &\vdots \\ |a_{nn}| &> |a_{n1}| + |a_{n2}| + \cdots + |a_{n,n-1}| \end{aligned}$$

Esto es: el valor absoluto de cada entrada diagonal $a_{11}, a_{22}, \ldots, a_{nn}$ es mayor que la suma de los valores absolutos de las entradas restantes en dicho renglón.

Teorema 2.9 Si un sistema de n ecuaciones lineales con n variables tiene una matriz coeficiente estrictamente diagonal dominante, entonces tiene una solución única y tanto el método de Jacobi como el de Gauss-Seidel convergen en ella.

Comentario ¡Atención! Este teorema es una implicación unidireccional. El hecho de que un sistema *no* sea estrictamente diagonal dominante *no* significa que los métodos iterativos diverjan. Pueden o no converger. (Vea los ejercicios 15-19.) De hecho, existen ejemplos en los que uno de los métodos converge y el otro diverge. Sin embargo, *si* alguno de los métodos converge, entonces debe converger en la solución: no puede converger en algún otro punto.

Teorema 2.10 Si el método de Jacobi o el de Gauss-Seidel convergen para un sistema de n ecuaciones lineales con n variables, entonces debe converger en la solución del sistema.

Demostración Se ilustrará la idea detrás de la demostración al bosquejarla para el caso del método de Jacobi, mediante el sistema de ecuaciones del ejemplo 2.37. La demostración general es similar.

Convergencia significa que "conforme aumentan las iteraciones, los valores de los iterados se acercan más a un valor límite". Esto significa que x_1 y x_2 convergen en r y s, respectivamente, como se muestra en la tabla 2.17.

Debe demostrar que $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} r \\ s \end{bmatrix}$ es la solución del sistema de ecuaciones. En otras palabras, en la $(k+1)$-ésima iteración, los valores de x_1 y x_2 deben permanecer igual que en la k-ésima iteración. Pero los cálculos producen $x_1 = (5 + x_2)/7 = (5 + s)/7$ y $x_2 = (7 + 3x_1)/5 = (7 + 3r)/5$. Por tanto,

Tabla 2.17

n	$\cdots$	k	$k+1$	$k+2$	$\cdots$
x_1	$\cdots$	r	r	r	$\cdots$
x_2	$\cdots$	s	s	s	$\cdots$

$$\frac{5+s}{7} = r \quad \text{y} \quad \frac{7+3r}{5} = s$$

Al reordenar, se ve que

$$\begin{aligned} 7r - s &= 5 \\ 3r - 5s &= -7 \end{aligned}$$

Por tanto, $x_1 = r$, $x_2 = s$ satisfacen las ecuaciones originales, como se requiere.

Por ahora puede preguntarse: si los métodos iterativos no siempre convergen en la solución, ¿qué tan buenos son? ¿Por qué no se usa la eliminación gaussiana? Primero, se vio que la eliminación gaussiana es sensible a errores de redondeo, y esta sensibilidad puede conducir a imprecisiones o incluso a respuestas enormemente equivocadas. Además, incluso si la eliminación gaussiana no se desviará, una solución no puede mejorarse una vez encontrada. Por ejemplo, si se usa la eliminación gaussiana para calcular una solución a dos lugares decimales, no hay forma de obtener la solución a cuatro lugares decimales, excepto comenzar nuevamente y trabajar con una precisión creciente.

En contraste, puede lograrse una precisión adicional con los métodos iterativos al simplemente hacer más iteraciones. Para sistemas grandes, en particular para aquellos con matrices de coeficientes dispersos, los métodos iterativos son mucho más rápidos que los métodos directos cuando se implementan en una computadora. En muchas aplicaciones, los sistemas que surgen son estrictamente diagonales dominantes, y por tanto está garantizado que los métodos iterativos converjan. El siguiente ejemplo ilustra una de tales aplicaciones.

Ejemplo 2.39

Suponga que se calienta el borde de una placa metálica a una temperatura constante, como se muestra en la figura 2.29.

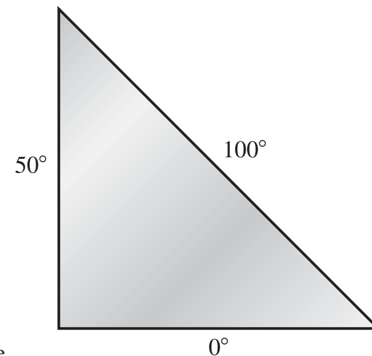


Figura 2.29
Una placa de metal caliente

Con el tiempo, la temperatura en los puntos interiores llegará al *equilibrio*, donde puede demostrarse que se cumple la siguiente propiedad:

La temperatura en cada punto interior P sobre una placa es el promedio de las temperaturas en la circunferencia de cualquier círculo con centro en P dentro de la placa (figura 2.30).

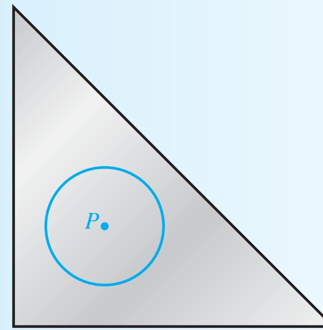


Figura 2.30

Aplicar esta propiedad en un ejemplo real requiere técnicas de cálculo. Como alternativa, puede aproximarse la situación al superponer a la placa una rejilla o malla, que tenga un número finito de puntos interiores, como se muestra en la figura 2.31.

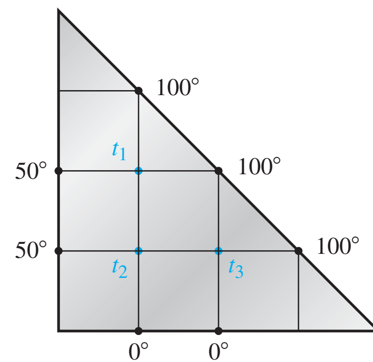


Figura 2.31
Versión discreta del problema de la placa caliente

A continuación se enuncia el análogo discreto de la propiedad promedio que gobierna las temperaturas de equilibrio:

La temperatura en cada punto interior P es el promedio de las temperaturas en los puntos adyacentes a P.

Para el ejemplo que se muestra en la figura 2.31, hay tres puntos interiores, y cada uno es adyacente a otros cuatro puntos. Sean t_1, t_2 y t_3 las temperaturas de equilibrio

de los puntos interiores, como se muestra. Entonces, por la propiedad de temperatura promedio, se tiene

$$
\begin{aligned}
t_1 &= \frac{100 + 100 + t_2 + 50}{4} \\
t_2 &= \frac{t_1 + t_3 + 0 + 50}{4} \\
t_3 &= \frac{100 + 100 + 0 + t_2}{4}
\end{aligned}
\tag{3}
$$

o

$$
\begin{aligned}
4t_1 - t_2 \quad\quad &= 250 \\
-t_1 + 4t_2 - t_3 &= 50 \\
- t_2 + 4t_3 &= 200
\end{aligned}
$$

Note que este sistema es estrictamente diagonal dominante. Observe también que las ecuaciones (3) están en la forma requerida por la iteración de Jacobi o Gauss-Seidel. Con una aproximación inicial de $t_1 = 0$, $t_2 = 0$, $t_3 = 0$, el método de Gauss-Seidel produce los siguientes iterados.

Iteración 1:

$$
\begin{aligned}
t_1 &= \frac{100 + 100 + 0 + 50}{4} = 62.5 \\
t_2 &= \frac{62.5 + 0 + 0 + 50}{4} = 28.125 \\
t_3 &= \frac{100 + 100 + 0 + 28.125}{4} = 57.031
\end{aligned}
$$

Iteración 2:

$$
\begin{aligned}
t_1 &= \frac{100 + 100 + 28.125 + 50}{4} = 69.531 \\
t_2 &= \frac{69.531 + 57.031 + 0 + 50}{4} = 44.141 \\
t_3 &= \frac{100 + 100 + 0 + 44.141}{4} = 61.035
\end{aligned}
$$

A continuación, se encuentran los iterados que se mencionan en la tabla 2.18. Se trabaja con una precisión de cinco dígitos significativos y se detiene cuando dos iterados sucesivos concuerdan dentro de 0.001 en todas la variables.



Por ende, las temperaturas de equilibrio en los puntos interiores son (con una precisión de 0.001): $t_1 = 74.108$, $t_2 = 46.430$, y $t_3 = 61.607$. (Compruebe estos cálculos.)

Al usar una retícula más fina (con más puntos interiores), puede obtener una información tan precisa como desee acerca de las temperaturas de equilibrio en varios puntos sobre la placa.

Tabla 2.18

n	0	1	2	3	⋯	7	8
t_1	0	62.500	69.531	73.535	⋯	74.107	74.107
t_2	0	28.125	44.141	46.143	⋯	46.429	46.429
t_3	0	57.031	61.035	61.536	⋯	61.607	61.607

Ejercicios 2.5

CAS

En los ejercicios 1-6, aplique el método de Jacobi al sistema dado. Tome el vector cero como la aproximación inicial y trabaje con una precisión de cuatro dígitos significativos hasta que dos iterados sucesivos concuerden dentro de 0.001 en cada variable. En cada caso, compare su respuesta con la solución exacta que encuentre usando cualquier método directo de su elección.

1. $\begin{aligned} 7x_1 - x_2 &= 6 \\ x_1 - 5x_2 &= -4 \end{aligned}$

2. $\begin{aligned} 2x_1 + x_2 &= 5 \\ x_1 - x_2 &= 1 \end{aligned}$

3. $\begin{aligned} 4.5x_1 - 0.5x_2 &= 1 \\ x_1 - 3.5x_2 &= -1 \end{aligned}$

4. $\begin{aligned} 20x_1 + x_2 - x_3 &= 17 \\ x_1 - 10x_2 + x_3 &= 13 \\ -x_1 + x_2 + 10x_3 &= 18 \end{aligned}$

5. $\begin{aligned} 3x_1 + x_2 &= 1 \\ x_1 + 4x_2 + x_3 &= 1 \\ x_2 + 3x_3 &= 1 \end{aligned}$

6. $\begin{aligned} 3x_1 - x_2 &= 1 \\ -x_1 + 3x_2 - x_3 &= 0 \\ -x_2 + 3x_3 - x_4 &= 1 \\ -x_3 + 3x_4 &= 1 \end{aligned}$

En los ejercicios 7-12, repita el ejercicio dado usando el método de Gauss-Seidel. Tome el vector cero como la aproximación inicial y trabaje con una precisión de cuatro dígitos significativos hasta que dos iterados sucesivos concuerden dentro de 0.001 en cada variable. Compare el número de iteraciones requeridas por los métodos de Jacobi y de Gauss-Seidel para alcanzar tal solución aproximada.

7. Ejercicio 1 **8.** Ejercicio 2

9. Ejercicio 3 **10.** Ejercicio 4

11. Ejercicio 5 **12.** Ejercicio 6

En los ejercicios 13 y 14, dibuje diagramas para ilustrar la convergencia del método de Gauss-Seidel con el sistema dado.

13. El sistema del ejercicio 1

14. El sistema del ejercicio 2

En los ejercicios 15 y 16, calcule los primeros cuatro iterados, con el vector cero como la aproximación inicial, para demostrar que el método de Gauss-Seidel diverge. Luego demuestre que las ecuaciones pueden reordenarse para dar una matriz de coeficientes estrictamente diagonal dominante y aplique el método de Gauss-Seidel para obtener una solución aproximada que sea precisa hasta dentro de 0.001.

15. $\begin{aligned} x_1 - 2x_2 &= 3 \\ 3x_1 + 2x_2 &= 1 \end{aligned}$

16. $\begin{aligned} x_1 - 4x_2 + 2x_3 &= 2 \\ 2x_2 + 4x_3 &= 1 \\ 6x_1 - x_2 - 2x_3 &= 1 \end{aligned}$

17. Dibuje un diagrama para ilustrar la divergencia del método de Gauss-Seidel en el ejercicio 15.

En los ejercicios 18 y 19, la matriz de coeficientes no es estrictamente diagonal dominante, ni pueden reordenarse las ecuaciones para hacerlas de ese modo. Sin embargo, tanto el método de Jacobi como el de Gauss-Seidel convergen de cualquier forma. Demuestre que esto es verdadero para el método de Gauss-Seidel, comience con el vector cero como la aproximación inicial y obtenga una solución que sea precisa hasta dentro de 0.01.

18. $\begin{aligned} -4x_1 + 5x_2 &= 14 \\ x_1 - 3x_2 &= -7 \end{aligned}$

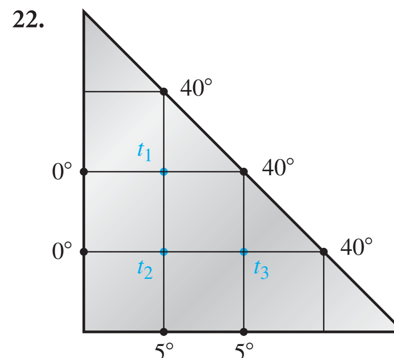
19. $\begin{aligned} 5x_1 - 2x_2 + 3x_3 &= -8 \\ x_1 + 4x_2 - 4x_3 &= 102 \\ -2x_1 - 2x_2 + 4x_3 &= -90 \end{aligned}$

20. Continúe con las iteraciones en el ejercicio 18 hasta obtener una solución que sea precisa hasta dentro de 0.001.

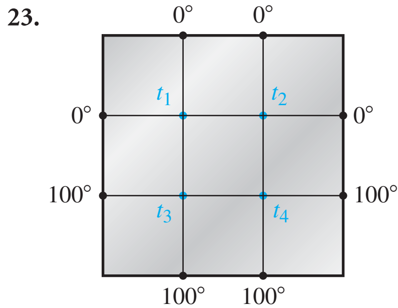
21. Continúe con las iteraciones en el ejercicio 19 hasta obtener una solución que sea precisa hasta dentro de 0.001.

En los ejercicios 22-24, la placa metálica tiene las temperaturas constantes que se muestran en sus fronteras. Encuentre la temperatura de equilibrio en cada uno de los puntos interiores indicados al establecer un sistema de ecuaciones lineales y aplicar el método de Jacobi o el de Gauss-Seidel. Obtenga una solución que sea precisa hasta dentro de 0.001.

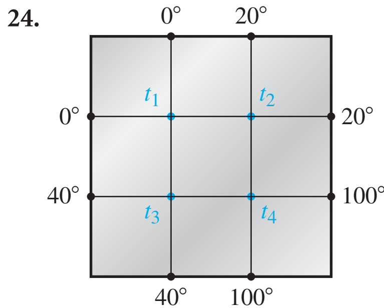
22.



23.

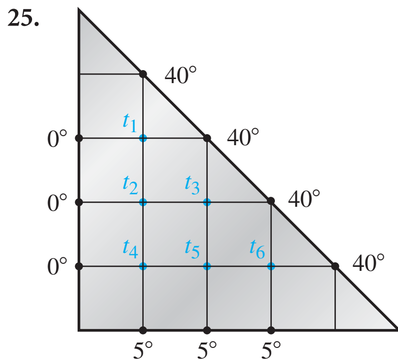


24.

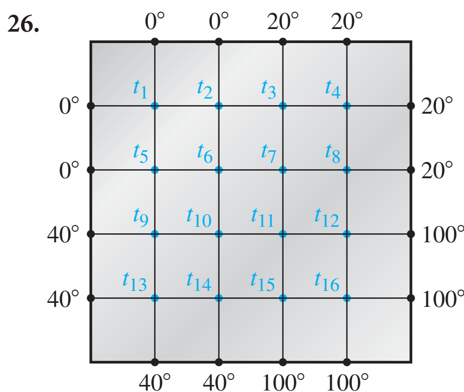


En los ejercicios 25 y 26, se refinan las retículas usadas en los ejercicios 22 y 24 para obtener información más precisa acerca de las temperaturas de equilibrio en los puntos interiores de las placas. Obtenga soluciones que sean precisas hasta dentro de 0.001, usando o el método de Jacobi o el de Gauss-Seidel.

25.



26.



Los ejercicios 27 y 28 demuestran que, a veces, si se tiene suerte, la forma de un problema iterativo puede permitir el uso de un poco de intuición para obtener una solución exacta.

27. Una tira estrecha de papel, de 1 unidad de longitud, se coloca a lo largo de una recta numérica, de modo que sus extremos están en 0 y 1. El papel se dobla a la mitad, con el extremo derecho sobre el izquierdo, de modo que sus extremos están ahora en 0 y $\frac{1}{2}$. A continuación, se dobla nuevamente a la mitad, esta vez con el extremo izquierdo sobre el derecho, de modo que sus extremos están en $\frac{1}{4}$ y $\frac{1}{2}$. La figura 2.32 muestra este proceso. Siga doblando el papel a la mitad, alternando derecha sobre izquierda e izquierda sobre derecha. Si pudiera continuar indefinidamente, es claro que los extremos del papel convergerían en un punto. Este punto es el que se quiere encontrar.

(a) Sea x_1 el extremo correspondiente al lado izquierdo del papel y x_2 el del extremo derecho. Elabore una tabla con los primeros seis valores de $[x_1, x_2]$ y grafique los correspondientes puntos sobre ejes coordenados x_1, x_2.

(b) Encuentre dos ecuaciones lineales de la forma $x_2 = ax_1 + b$ y $x_1 = cx_2 + d$ que determinen los nuevos valores de los puntos finales de cada iteración. Dibuje las rectas correspondientes en sus ejes coordenados y demuestre que este diagrama resultaría de aplicar el método de Gauss-Seidel al sistema de ecuaciones lineales que encontró. (Su diagrama debe parecerse al de la figura 2.27, de la página 132.)

(c) Cambie a representación decimal y siga aplicando el método de Gauss-Seidel para aproximar el punto donde convergen los extremos del papel hasta dentro de 0.001 de precisión.

(d) Resuelva el sistema de ecuaciones exactamente y compare sus respuestas.

28. Una hormiga está parada sobre una recta numérica en el punto A. Camina a la mitad hacia el punto B y da vuelta. Luego camina la mitad de vuelta al punto A, da vuelta nuevamente y camina a la mitad hacia el punto B. Y sigue haciendo esto indefinidamente. Sea que el punto A está en 0 y el punto B en 1. El paseo de la hormiga está constituido por una sucesión de segmentos de recta que se traslapan. Sea x_1 el registro de las posiciones de los puntos finales izquierdos de dichos segmentos y x_2 sus puntos finales derechos. (Por tanto, comience con $x_1 = 0$ y $x_2 = \frac{1}{2}$. Luego se tiene $x_1 = \frac{1}{4}$ y $x_2 = \frac{1}{2}$, etcétera.) La figura 2.33 muestra el inicio del paseo de la hormiga.

(a) Elabore una tabla con los primeros seis valores de $[x_1, x_2]$ y grafique los puntos correspondientes en ejes coordenados x_1, x_2.

(b) Encuentre dos ecuaciones lineales de la forma $x_2 = ax_1 + b$ y $x_1 = cx_2 + d$ que determinen los nuevos valores de

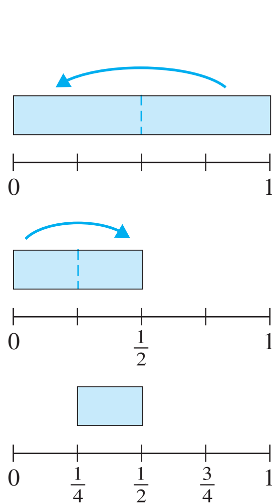


Figura 2.32
Doblado de una tira de papel

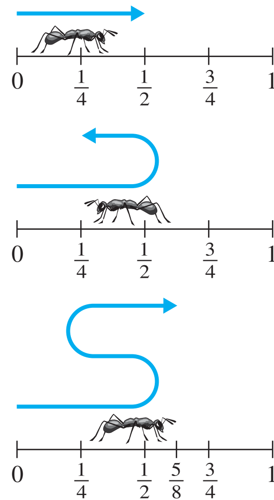


Figura 2.33
El paseo de la hormiga

los puntos finales en cada iteración. Dibuje las rectas correspondientes sobre sus ejes coordenados y demuestre que este diagrama resultaría de aplicar el método de Gauss-Seidel al sistema de ecuaciones lineales que encontró. (Su diagrama debe parecerse al de la figura 2.27 de la página 132.)

(c) Cambie a representación decimal y siga aplicando el método de Gauss-Seidel para aproximar los valores en los cuales convergen x_1 y x_2 hasta una precisión dentro de 0.001.

(d) Resuelva el sistema de ecuaciones exactamente y compare sus respuestas. Interprete sus resultados.

Repaso del capítulo

Definiciones y conceptos clave

conjunto generador, 96
convergencia, 131-132
divergencia, 133
ecuación lineal, 63
eliminación de Gauss-Jordan, 79
eliminación gaussiana, 74
forma escalonada por renglones, 71
forma escalonada reducida por renglones, 79
generador de un conjunto de vectores, 96
iterado, 131
matrices equivalentes por renglones, 74
matriz aumentada, 67
matriz coeficiente, 70
método de Gauss-Seidel, 130
método de Jacobi, 130
operaciones elementales con renglones, 72
pivote, 72
rank de una matriz, 78
sistema de ecuaciones lineales, 64
sustitución hacia atrás, 67
sistema consistente, 66
sistema homogéneo, 82
sistema inconsistente, 66
teorema del rank, 78
variable libre, 77
vectores linealmente dependientes, 99
vectores linealmente independientes, 99
variable pivote (1 pivote), 77-79

Preguntas de repaso

1. Marque cada uno de los siguientes enunciados como verdadero o falso:

(a) Todo sistema de ecuaciones lineales tiene una solución.

(b) Todo sistema homogéneo de ecuaciones lineales tiene una solución.

(c) Si un sistema de ecuaciones lineales tiene más variables que ecuaciones, entonces tiene un número infinito de soluciones.

(d) Si un sistema de ecuaciones lineales tiene más ecuaciones que variables, entonces no tiene solución.

(e) Determinar si $\mathbf{b}$ está en gen$(\mathbf{a}_1, \ldots, \mathbf{a}_n)$ es equivalente a determinar si el sistema $[A \mid \mathbf{b}]$ es consistente, donde $A = [\mathbf{a}_1 \cdots \mathbf{a}_n]$.

(f) En $\mathbb{R}^3$, gen$(\mathbf{u}, \mathbf{v})$ siempre es un plano que pasa por el origen.

(g) En $\mathbb{R}^3$, si los vectores $\mathbf{u}$ y $\mathbf{v}$ distintos de cero no son paralelos, entonces son linealmente independientes.

(h) En $\mathbb{R}^3$, si un conjunto de vectores puede dibujarse punta a origen, uno después del otro, de modo que se forme una trayectoria cerrada (polígono), entonces los vectores son linealmente dependientes.

(i) Si un conjunto de vectores tiene la propiedad de que ningún par de vectores en el conjunto son múltiplos escalares mutuos, entonces el conjunto de vectores es linealmente independiente.

(j) Si en un conjunto de vectores existen más vectores que el número de elementos en cada vector, entonces el conjunto de vectores es linealmente dependiente.

2. Encuentre el rank de la matriz

$$\begin{bmatrix} 1 & -2 & 0 & 3 & 2 \\ 3 & -1 & 1 & 3 & 4 \\ 3 & 4 & 2 & -3 & 2 \\ 0 & -5 & -1 & 6 & 2 \end{bmatrix}.$$

3. Resuelva el sistema lineal

$$\begin{aligned} x + y - 2z &= 4 \\ x + 3y - z &= 7 \\ 2x + y - 5z &= 7 \end{aligned}$$

4. Resuelva el sistema lineal

$$\begin{aligned} 3w + 8x - 18y + z &= 35 \\ w + 2x - 4y &= 11 \\ w + 3x - 7y + z &= 10 \end{aligned}$$

5. Resuelva el sistema lineal

$$\begin{aligned} 2x + 3y &= 4 \\ x + 2y &= 3 \end{aligned}$$

sobre $\mathbb{Z}_7$.

6. Resuelva el sistema lineal

$$\begin{aligned} 3x + 2y &= 1 \\ x + 4y &= 2 \end{aligned}$$

sobre $\mathbb{Z}_5$.

7. ¿Para qué valor(es) de k es inconsistente el sistema lineal con matriz aumentada $\left[\begin{array}{cc|c} k & 2 & 1 \\ 1 & 2k & 1 \end{array}\right]$?

8. Encuentre ecuaciones paramétricas para la recta de intersección de los planos $x + 2y + 3z = 4$ y $5x + 6y + 7z = 8$.

9. Encuentre el punto de intersección de las siguientes rectas, si existe.

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + s\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \\ -4 \end{bmatrix} + t\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

10. Determine si $\begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}$ está en el generador de $\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$.

11. Encuentre la ecuación general del plano generado por $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$.

12. Determine si $\begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -2 \end{bmatrix}, \begin{bmatrix} 3 \\ 9 \\ -2 \end{bmatrix}$ son linealmente independientes.

13. Determine si $\mathbb{R}^3 = \text{gen}(\mathbf{u}, \mathbf{v}, \mathbf{w})$ si:

(a) $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(b) $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$

14. Sean $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ vectores linealmente independientes en $\mathbb{R}^3$, y sea $A = [\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{a}_3]$. ¿Cuál de los siguientes enunciados es verdadero?

(a) La forma escalonada reducida por renglones de A es I_3.

(b) El rank de A es 3.

(c) El sistema $[A \mid \mathbf{b}]$ tiene una solución única para cualquier vector $\mathbf{b}$ en $\mathbb{R}^3$.

(d) (a), (b) y (c) son verdaderos.

(e) (a) y (b) son verdaderos, pero (c) no lo es.

15. Sean $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ vectores linealmente dependientes en $\mathbb{R}^3$, no todos cero, y sea $A = [\mathbf{a}_1 \ \ \mathbf{a}_2 \ \ \mathbf{a}_3]$. ¿Cuáles son los posibles valores del rank de A?

16. ¿Cuál es el rank máximo de una matriz de 5×3? ¿Cuál es el rango mínimo de una matriz de 5×3?

17. Demuestre que, si $\mathbf{u}$ y $\mathbf{v}$ son vectores linealmente independientes, entonces también lo son $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$.

18. Demuestre que $\text{gen}(\mathbf{u}, \mathbf{v}) = \text{span}(\mathbf{u}, \mathbf{u} + \mathbf{v})$ para cualesquiera vectores $\mathbf{u}$ y $\mathbf{v}$.

19. Con la finalidad de que un sistema lineal con matriz aumentada $[A \mid \mathbf{b}]$ sea consistente, ¿qué debe ser cierto acerca de los ranks de A y $[A \mid \mathbf{b}]$?

20. ¿Las matrices $\begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & -1 \\ -1 & 4 & 1 \end{bmatrix}$ y $\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & 3 \end{bmatrix}$ son equivalentes por renglones? ¿Por qué sí o por qué no?

3 Matrices

Nosotros [Halmos y Kaplansky] compartimos una filosofía acerca del álgebra lineal: pensamos libres de bases, escribimos libres de bases, pero a la hora de la verdad, cerramos la puerta de la oficina y calculamos con matrices como enajenados.

—Irving Kaplansky
en *Paul Halmos: Celebrating 50 Years of Mathematics*
J. H. Ewing y F. W. Gehring, eds.,
Springer-Verlag, 1991, p. 88

3.0 Introducción: matrices en acción

En este capítulo se estudiarán las matrices por derecho propio. Ya se usaron matrices, en forma de matrices aumentadas, para registrar información acerca de los sistemas de ecuaciones lineales y para ayudar a agilizar los cálculos que los involucran. Ahora verá que las matrices tienen propiedades algebraicas propias, lo que permite realizar cálculos con ellas, sujetas a las reglas del álgebra matricial. Más aún, observará que las matrices no son objetos estáticos que registran información y datos; en vez de ello, representan ciertos tipos de funciones que "actúan" sobre los vectores y los transforman en otros vectores. Tales "transformaciones matriciales" comenzarán a tener un papel clave en el estudio del álgebra lineal y arrojarán nueva luz en lo que ya aprendió acerca de los vectores y los sistemas de ecuaciones lineales. Además, las matrices surgen en muchas formas diferentes a las de las matrices aumentadas; al final de este capítulo se explorarán algunas de las muchas aplicaciones de las matrices.

En esta sección se considerarán algunos ejemplos simples para ilustrar cómo las matrices pueden transformar a los vectores. En el proceso tendrá un primer vistazo de la "aritmética matricial".

Considere las ecuaciones

$$\begin{aligned} y_1 &= x_1 + 2x_2 \\ y_2 &= \phantom{x_1 + {}} 3x_2 \end{aligned} \qquad (1)$$

Puede considerar que estas ecuaciones describen una transformación del vector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ en el vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$. Si se denota con F la matriz de coeficientes del lado derecho, entonces $F = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$ y puede reescribir la transformación como

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

o, de manera más sucinta, $\mathbf{y} = F\mathbf{x}$. (Piense en esta expresión como el análogo de la notación funcional $y = f(x)$ a la que está acostumbrado: $\mathbf{x}$ aquí es la "variable" independiente, $\mathbf{y}$ es la "variable" dependiente y F es el nombre de la "función".)

Por tanto, si $\mathbf{x} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, entonces las ecuaciones (1) producen

$$\begin{aligned} y_1 &= -2 + 2 \cdot 1 = 0 \\ y_2 &= \quad\;\; 3 \cdot 1 = 3 \end{aligned} \quad \text{o} \quad \mathbf{y} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Esta expresión puede escribirse como $\begin{bmatrix} 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Problema 1 Calcule $F\mathbf{x}$ para los siguientes vectores $\mathbf{x}$:

(a) $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ (b) $\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ (c) $\mathbf{x} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$ (d) $\mathbf{x} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

Problema 2 Las puntas de los cuatro vectores $\mathbf{x}$ del problema 1 se ubican en las cuatro esquinas de un cuadrado en el plano x_1x_2. Dibuje este cuadrado y marque sus esquinas A, B, C y D, que corresponden a los incisos (a), (b), (c) y (d) del problema 1.

En ejes coordenados separados (marcados y_1 y y_2), dibuje los cuatro puntos determinados por $F\mathbf{x}$ en el problema 1. Marque estos puntos A', B', C' y D'. Haga la suposición (razonable) de que el segmento de recta $\overline{AB}$, se transforma en el segmento de recta $\overline{A'B'}$ y del mismo modo para los otros tres lados del cuadrado $ABCD$. ¿Qué figura geométrica se representa mediante $A'B'C'D'$?

Problema 3 El centro del cuadrado $ABCD$ es el origen $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. ¿Cuál es el centro de $A'B'C'D'$? ¿Qué cálculo algebraico confirma esto?

Ahora considere las ecuaciones

$$\begin{aligned} z_1 &= \quad y_1 - y_2 \\ z_2 &= -2y_1 \end{aligned} \qquad (2)$$

que transforman un vector $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ en el vector $\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$. Puede abreviar esta transformación como $\mathbf{z} = G\mathbf{y}$, de donde

$$G = \begin{bmatrix} 1 & -1 \\ -2 & 0 \end{bmatrix}$$

Problema 4 Va a descubrir cómo G transforma a la figura $A'B'C'D'$. Calcule $G\mathbf{y}$ para cada uno de los cuatro vectores $\mathbf{y}$ que calculó en el problema 1. [Esto es: calcule $\mathbf{z} = G(F\mathbf{x})$. Puede reconocer esta expresión como análoga a la composición de funciones con la que está familiarizado.] Llame a los puntos correspondientes A'', B'', C'' y D'', y bosqueje la figura $A''B''C''D''$ sobre los ejes coordenados z_1z_2.

Problema 5 Al sustituir las ecuaciones (1) en las ecuaciones (2), se obtienen ecuaciones para z_1 y z_2 en términos de x_1 y x_2. Si denota con H la matriz de estas ecuaciones, entonces se tiene $\mathbf{z} = H\mathbf{x}$. Dado que también se tiene $\mathbf{z} = GF\mathbf{x}$, es razonable escribir

$$H = GF$$

¿Puede ver cómo las entradas de H se relacionan con las entradas de F y G?

Problema 6 Repita el proceso anterior en forma inversa: primero transforme el cuadrado $ABCD$, usando G, para obtener la figura $A^*B^*C^*D^*$. Luego transforme la figura resultante, con F, para obtener $A^{**}B^{**}C^{**}D^{**}$. [*Nota:* no se preocupe aquí por las "variables" $\mathbf{x}$, $\mathbf{y}$ y $\mathbf{z}$. Simplemente sustituya las coordenadas de A, B, C y D en las ecuaciones (2) y luego sustituya los resultados en las ecuaciones (1).] ¿$A^{**}B^{**}C^{**}D^{**}$ y

$A''B''C''D''$ son iguales? ¿Qué le dice esto acerca del orden en que se realizan las transformaciones F y G?

Problema 7 Repita el problema 5 con las matrices generales

$$F = \begin{bmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{bmatrix}, \quad G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad \text{y} \quad H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}$$

Esto es: si las ecuaciones (1) y las ecuaciones (2) tienen coeficientes como especifican F y G, encuentre las entradas de H en términos de las entradas de F y G. El resultado será una fórmula para el "producto" $H = GF$.

Problema 8 Repita los problemas 1-6 con las siguientes matrices. (Su fórmula del problema 7 le ayudará a acelerar los cálculos algebraicos). Note cualquier similitud o diferencia que considere significativa.

(a) $F = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, G = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ (b) $F = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, G = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

(c) $F = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, G = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$ (d) $F = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}, G = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$

3.1 Operaciones con matrices

Aunque ya se ha topado con las matrices, se comienza por enunciar una definición formal y registrar algunos hechos para futura referencia.

Definición Una ***matriz*** es un arreglo rectangular de números llamados ***entradas***, o ***elementos***, de la matriz.

Aunque los números por lo general se elegirán del conjunto $\mathbb{R}$ de números reales, también pueden tomarse del conjunto $\mathbb{C}$ de números complejos o de $\mathbb{Z}_p$ donde p es primo.

Los siguientes son ejemplos de matrices:

$$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} \sqrt{5} & -1 & 0 \\ 2 & \pi & \frac{1}{2} \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 4 \\ 17 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 5.1 & 1.2 & -1 \\ 6.9 & 0 & 4.4 \\ -7.3 & 9 & 8.5 \end{bmatrix}, \quad [7]$$

El ***tamaño*** de una matriz es una descripción de los números de renglones y columnas que tiene. Una matriz se llama de $m \times n$ (dígase "m por n") si tiene m renglones y n columnas. Por tanto, los ejemplos anteriores son matrices de tamaños 2×2, 2×3, 3×1, 1×4, 3×3 y 1×1, respectivamente. Una matriz de $1 \times n$ se llama ***matriz renglón*** (o ***vector renglón***), y una matriz de $n \times 1$ se llama ***matriz columna*** (o ***vector columna***).

Técnicamente, hay distinción entre matrices renglón/columna y vectores, pero aquí no se profundizará en dicha distinción. Sin embargo, *sí* se distinguirá entre matrices/vectores *renglón* y matrices/vectores *columna*. Esta distinción es importante, al menos, para cálculos algebraicos, como se demostrará.

Se usa la notación de *doble subíndice* para referirse a las entradas de una matriz A. La entrada de A en el renglón i y la columna j se denota mediante a_{ij}. Por tanto, si

$$A = \begin{bmatrix} 3 & 9 & -1 \\ 0 & 5 & 4 \end{bmatrix}$$

entonces $a_{13} = -1$ y $a_{22} = 5$. (En ocasiones se usa la notación A_{ij} de manera intercambiable con a_{ij}.) Por tanto, de manera compacta, una matriz A se puede denotar mediante $[a_{ij}]$ (o $[a_{ij}]_{m \times n}$ si es importante especificar el tamaño de A, aunque el tamaño por lo general será claro a partir del contexto).

Con esta notación, una matriz general A de $m \times n$ tiene la forma

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Si las columnas de A son los vectores $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$, entonces A se puede representar como

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]$$

Si los renglones de A son $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_m$, entonces A se puede presentar como

$$A = \begin{bmatrix} \mathbf{A}_1 \\ \mathbf{A}_2 \\ \vdots \\ \mathbf{A}_m \end{bmatrix}$$

Las ***entradas diagonales*** de A son $a_{11}, a_{22}, a_{33}, \ldots,$ y si $m = n$ (esto es, si A tiene el mismo número de renglones que de columnas), entonces A es una ***matriz cuadrada***. Una matriz cuadrada cuyas entradas no diagonales sean todas cero es una ***matriz diagonal***. Una matriz diagonal cuyas entradas diagonales sean todas iguales es una ***matriz escalar***. Si el escalar en la diagonal es 1, la matriz escalar es una ***matriz identidad***.

Por ejemplo, sean

$$A = \begin{bmatrix} 2 & 5 & 0 \\ -1 & 4 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 \\ 4 & 5 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Las entradas diagonales de A son 2 y 4, pero A no es cuadrada; B es una matriz cuadrada de tamaño 2×2, con entradas diagonales 3 y 5; C es una matriz diagonal; D es una matriz identidad de 3×3. La matriz identidad $n \times n$ se denota I_n (o simplemente I si su tamaño se entiende).

Dado que las matrices pueden verse como generalizaciones de vectores (y, de hecho, las matrices pueden y deben considerarse como constituidas tanto por vectores renglón como columna), muchas de las convenciones y operaciones para vectores se realizan sobre las matrices (en una forma obvia).

Dos matrices son ***iguales*** si tienen el mismo tamaño y si sus entradas correspondientes son iguales. Por tanto, si $A = [a_{ij}]_{m \times n}$ y $B = [b_{ij}]_{r \times s}$, entonces $A = B$ si y sólo si $m = r$ y $n = s$ y $a_{ij} = b_{ij}$ para todo i y j.

Ejemplo 3.1

Considere las matrices

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 5 & 3 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 2 & 0 & x \\ 5 & 3 & y \end{bmatrix}$$

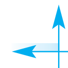
Ni A ni B pueden ser iguales a C (sin importar cuáles sean los valores de x y y), pues A y B son matrices de 2×2 y C es de 2×3. Sin embargo, $A = B$ si y sólo si $a = 2$, $b = 0$, $c = 5$ y $d = 3$.

Ejemplo 3.2

Considere las matrices

$$R = [1 \quad 4 \quad 3] \quad \text{y} \quad C = \begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$$

A pesar del hecho de que R y C tienen las mismas entradas en el mismo orden, $R \neq C$, pues R es 1×3 y C es 3×1. (Si R y C se leen en voz alta, suenan igual: "uno, cuatro, tres".) Por tanto, la distinción entre matrices/vectores renglón y matrices/vectores columna es importante.



Suma de matrices y multiplicación por un escalar

Para generalizar la suma de vectores, se define la suma de matrices *por componentes*. Si $A = [a_{ij}]$ y $B = [b_{ij}]$ son matrices de $m \times n$, su ***suma*** $A + B$ es la matriz de $m \times n$ que se obtiene al sumar las entradas correspondientes. Por tanto,

$$A + B = [a_{ij} + b_{ij}]$$

[Igualmente se podría definir $A + B$ en términos de suma vectorial al especificar que cada columna (o renglón) de $A + B$ es la suma de las correspondientes columnas (o renglones) de A y B.] Si A y B no son del mismo tamaño, entonces $A + B$ no está definida.

Ejemplo 3.3

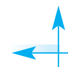
Sean

$$A = \begin{bmatrix} 1 & 4 & 0 \\ -2 & 6 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} -3 & 1 & -1 \\ 3 & 0 & 2 \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$$

Entonces

$$A + B = \begin{bmatrix} -2 & 5 & -1 \\ 1 & 6 & 7 \end{bmatrix}$$

pero ni $A + C$ ni $B + C$ están definidas.



La definición por componentes de la multiplicación por un escalar no será sorpresa. Si A es una matriz $m \times n$ y c es un escalar, entonces el ***múltiplo escalar*** cA es la matriz $m \times n$ que se obtiene al multiplicar cada entrada de A por c. De manera más formal, se tiene

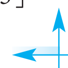
$$cA = c[a_{ij}] = [ca_{ij}]$$

[En términos de vectores, podría estipularse de manera equivalente que cada columna (o renglón) de cA es c veces la correspondiente columna (o renglón) de A.]

Ejemplo 3.4

Para la matriz A del ejemplo 3.3,

$$2A = \begin{bmatrix} 2 & 8 & 0 \\ -4 & 12 & 10 \end{bmatrix}, \quad \tfrac{1}{2}A = \begin{bmatrix} \frac{1}{2} & 2 & 0 \\ -1 & 3 & \frac{5}{2} \end{bmatrix} \quad \text{y} \quad (-1)A = \begin{bmatrix} -1 & -4 & 0 \\ 2 & -6 & -5 \end{bmatrix}$$



La matriz $(-1)A$ se escribe como $-A$ y se llama ***negativo*** de A. Como con los vectores, puede usar este hecho para definir la ***diferencia*** de dos matrices: si A y B tienen el mismo tamaño, entonces

$$A - B = A + (-B)$$

Ejemplo 3.5

Para las matrices A y B del ejemplo 3.3,

$$A - B = \begin{bmatrix} 1 & 4 & 0 \\ -2 & 6 & 5 \end{bmatrix} - \begin{bmatrix} -3 & 1 & -1 \\ 3 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 3 & 1 \\ -5 & 6 & 3 \end{bmatrix}$$

Una matriz cuyas entradas sean todas cero se llama ***matriz cero*** y se denota O (u $O_{m\times n}$ si es importante especificar su tamaño). Debe ser claro que si A es cualquier matriz y O es la matriz cero del mismo tamaño, entonces

$$A + O = A = O + A$$

y

$$A - A = O = -A + A$$

Multiplicación de matrices

En ocasiones, los matemáticos son como el Humpty Dumpty de Lewis Carroll: "Cuando *yo* uso una palabra", dijo Humpty Dumpty, "significa justo lo que quiero que signifique, ni más ni menos" (de *Through the Looking Glass*).

La introducción de la sección 3.0 sugirió que hay un "producto" de matrices que es análoga a la composición de funciones. Ahora se hará más precisa esta noción. La definición que está a punto de proporcionarse generaliza lo que debió descubrir en los problemas 5 y 7 de la sección 3.0. A diferencia de las definiciones de suma de matrices y multiplicación por un escalar, la definición del producto de dos matrices *no* es por componentes. Desde luego, no hay nada que impida definir un producto de matrices en forma de componentes; por desgracia, tal definición tiene pocas aplicaciones y no es tan "natural" como la que ahora se propondrá.

Definición Si A es una matriz de $m \times n$ y B es una matriz de $n \times r$, entonces el ***producto*** $C = AB$ es una matriz de $m \times r$. La entrada (i, j) del producto se calcula del modo siguiente:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$$

Comentarios

- Note que A y B no tienen que ser del mismo tamaño. Sin embargo, el número de *columnas* de A debe ser igual al número de *renglones* de B. Si escribe en orden los tamaños de A, B y AB, podrá ver de inmediato si se satisface este requisito. Más aún, puede predecir el tamaño del producto antes de hacer cálculo alguno, pues el número de *renglones* de AB es igual que el número de renglones de A, mientras que el número de *columnas* de AB es igual que el número de columnas de B, como se muestra a continuación:

$$\begin{array}{ccccc} A & & B & = & AB \\ m\times n & & n\times r & & m\times r \end{array}$$

Igual

Tamaño de AB

- La fórmula para las entradas del producto parece un producto punto, y de hecho lo es. Se dice que la entrada (i, j) de la matriz AB es el producto punto del i-ésimo renglón de A y la j-ésima columna de B:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1r} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2r} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nr} \end{bmatrix}$$

Note que en la expresión $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}$, los "subíndices exteriores" de cada término ab en la suma siempre son i y j, mientras que los "subíndices interiores" siempre concuerdan y aumentan de 1 a n. Este patrón se ve claramente si escribe c_{ij} usando la notación de suma:

$$c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Ejemplo 3.6

Calcule AB si

$$A = \begin{bmatrix} 1 & 3 & -1 \\ -2 & -1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -4 & 0 & 3 & -1 \\ 5 & -2 & -1 & 1 \\ -1 & 2 & 0 & 6 \end{bmatrix}$$

Solución Dado que A es de 2×3 y B es de 3×4, el producto AB está definido y será una matriz 2×4. El primer renglón del producto $C = AB$ se efectúa al tomar el producto punto del primer renglón de A con cada una de las columnas de B por turno. En consecuencia,

$$\begin{aligned} c_{11} &= 1(-4) + 3(5) + (-1)(-1) = 12 \\ c_{12} &= 1(0) + 3(-2) + (-1)(2) = -8 \\ c_{13} &= 1(3) + 3(-1) + (-1)(0) = 0 \\ c_{14} &= 1(-1) + 3(1) + (-1)(6) = -4 \end{aligned}$$

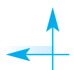
El segundo renglón de C se calcula al realizar el producto punto del segundo renglón de A con cada una de las columnas de B por turno:

$$\begin{aligned} c_{21} &= (-2)(-4) + (-1)(5) + (1)(-1) = 2 \\ c_{22} &= (-2)(0) + (-1)(-2) + (1)(2) = 4 \\ c_{23} &= (-2)(3) + (-1)(-1) + (1)(0) = -5 \\ c_{24} &= (-2)(-1) + (-1)(1) + (1)(6) = 7 \end{aligned}$$

Por tanto, la matriz producto está dada por

$$AB = \begin{bmatrix} 12 & -8 & 0 & -4 \\ 2 & 4 & -5 & 7 \end{bmatrix}$$

(Con un poco de práctica, deberá hacer estos cálculos mentalmente sin escribir todos los detalles como se hizo aquí. Para ejemplos más complicados, es preferible una calculadora con capacidades matriciales o un sistema algebraico de cómputo.)



Antes de avanzar más, se considerarán dos ejemplos que justifican la definición elegida de multiplicación de matrices.

Ejemplo 3.7

Ann y Bert planean ir a comprar frutas para la siguiente semana. Cada uno quiere comprar algunas manzanas, naranjas y uvas, pero en diferentes cantidades. La tabla 3.1 menciona lo que quieren comprar. Hay dos mercados de frutas cercanos, Sam's y Theo's, y sus precios se proporcionan en la tabla 3.2. ¿Cuánto costará a Ann y Bert hacer sus compras en cada uno de los dos mercados?

Tabla 3.1

	Manzanas	Uvas	Naranjas
Ann	6	3	10
Bert	4	8	5

Tabla 3.2

	Sam's	Theo's
Manzana	\$0.10	\$0.15
Uva	\$0.40	\$0.30
Naranja	\$0.10	\$0.20

Solución Si Ann compra en Sam's, gastará

$$6(0.10) + 3(0.40) + 10(0.10) = \$2.80$$

Si compra en Theo's, gastará

$$6(0.15) + 3(0.30) + 10(0.20) = \$3.80$$

Bert gastará

$$4(0.10) + 8(0.40) + 5(0.10) = \$4.10$$

en Sam's y

$$4(0.15) + 8(0.30) + 5(0.20) = \$4.00$$

en Theo's. (Presumiblemente, Ann comprará en Sam's mientras que Bert irá a Theo's.)

La "forma producto punto" de estos cálculos sugieren que la multiplicación de matrices funciona aquí. Si la información dada se organiza en una matriz demanda D y una matriz precio P, se tiene

$$D = \begin{bmatrix} 6 & 3 & 10 \\ 4 & 8 & 5 \end{bmatrix} \quad \text{y} \quad P = \begin{bmatrix} 0.10 & 0.15 \\ 0.40 & 0.30 \\ 0.10 & 0.20 \end{bmatrix}$$

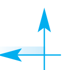
Los cálculos anteriores son equivalentes a calcular el producto

$$DP = \begin{bmatrix} 6 & 3 & 10 \\ 4 & 8 & 5 \end{bmatrix} \begin{bmatrix} 0.10 & 0.15 \\ 0.40 & 0.30 \\ 0.10 & 0.20 \end{bmatrix} = \begin{bmatrix} 2.80 & 3.80 \\ 4.10 & 4.00 \end{bmatrix}$$

Por tanto, la matriz producto DP dice cuánto costarán las compras de cada persona en cada tienda (tabla 3.3).

Tabla 3.3

	Sam's	Theo's
Ann	\$2.80	\$3.80
Bert	\$4.10	\$4.00



Ejemplo 3.8

Considere el sistema lineal

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 5 \\ -x_1 + 3x_2 + x_3 &= 1 \\ 2x_1 - x_2 + 4x_3 &= 14 \end{aligned} \qquad (1)$$

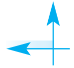
Observe que el lado izquierdo surge del producto matricial

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

de modo que el sistema (1) puede escribirse como

$$\begin{bmatrix} 1 & -2 & 3 \\ -1 & 3 & 1 \\ 2 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ 14 \end{bmatrix}$$

o $A\mathbf{x} = \mathbf{b}$, donde A es la matriz de coeficientes, $\mathbf{x}$ es el vector (columna) de variables y $\mathbf{b}$ es el vector (columna) de términos constantes.



El lector no debe tener dificultades para ver que *todo* sistema lineal puede escribirse en la forma $A\mathbf{x} = \mathbf{b}$. De hecho, la notación $[A \mid \mathbf{b}]$ para la matriz aumentada de un sistema lineal sólo es una abreviatura para la ecuación matricial $A\mathbf{x} = \mathbf{b}$. Esta forma probará ser tremendamente útil para expresar un sistema de ecuaciones lineales y se le explotará con frecuencia a partir de ahora.

Al combinar esta intuición con el Teorema 2.4, se ve que $A\mathbf{x} = \mathbf{b}$ tiene una solución si y sólo si $\mathbf{b}$ es una combinación lineal de las columnas de A.

Hay otro factor acerca de las operaciones matriciales que también resultará ser bastante útil: la multiplicación de una matriz por un vector unitario estándar puede usarse para "obtener" o "reproducir" una columna o un renglón de una matriz. Sea $A = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix}$ y considere los productos $A\mathbf{e}_3$ y $\mathbf{e}_2A$, con los vectores unitarios $\mathbf{e}_3$ y $\mathbf{e}_2$ elegidos de modo que los productos tengan sentido. Por tanto

$$A\mathbf{e}_3 = \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{y} \quad \mathbf{e}_2A = [0 \quad 1] \begin{bmatrix} 4 & 2 & 1 \\ 0 & 5 & -1 \end{bmatrix}$$

$$= [0 \quad 5 \quad -1]$$

Note que $A\mathbf{e}_3$ produce la tercera columna de A y $\mathbf{e}_2A$ produce el segundo renglón de A. El resultado general se registra como teorema.

Teorema 3.1

Sea A una matriz $n \times m$, e_i, un vector unitario estándar $1 \times m$ y $\mathbf{e}_j$ un vector unitario $n \times 1$ estándar. Entonces

a. $\mathbf{e}_i$ es el i-ésimo renglón de A y
b. $A\mathbf{e}_j$ es la j-ésima columna de A.

Demostración Se demuestra (b) y la demostración de (a) se deja como ejercicio 41. Si $\mathbf{a}_1, \ldots, \mathbf{a}_n$ son las columnas de A, entonces el producto $A\mathbf{e}_j$ puede escribirse

$$A\mathbf{e}_j = 0\mathbf{a}_1 + 0\mathbf{a}_2 + \cdots + 1\mathbf{a}_j + \cdots + 0\mathbf{a}_n = \mathbf{a}_j$$

También podría demostrar (b) mediante cálculo directo:

$$A\mathbf{e}_j = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

pues el 1 en $\mathbf{e}_j$ es la j-ésima entrada.

Matrices particionadas

Con frecuencia será conveniente considerar que una matriz está compuesta de algunas ***submatrices*** más pequeñas. Al introducir rectas verticales y horizontales en una matriz, se le puede ***partir*** en ***bloques***. Hay una forma natural de partir muchas matrices, en particular aquellas que surgen en ciertas aplicaciones. Por ejemplo, considere la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{bmatrix}$$

Parece natural partir A como

$$\left[\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] = \begin{bmatrix} I & B \\ O & C \end{bmatrix}$$

donde I es la matriz identidad de 3×3, B es de 3×2, O es la matriz cero de 2×3 y C es de 2×2. De esta forma, puede ver a A como una matriz de 2×2 cuyas entradas son ellas mismas matrices.

Cuando se multiplican matrices, con frecuencia se gana una ventaja al verlas como matrices particionadas. Esto frecuentemente no sólo revela las estructuras subyacentes, sino que usualmente acelera los cálculos, en especial cuando las matrices son grandes y tienen muchos bloques de ceros. Es evidente que la multiplicación de matrices particionadas es igual que la multiplicación ordinaria de matrices.

Comience por considerar algunos casos especiales de matrices particionadas. Cada una da origen a una forma diferente de ver el producto de dos matrices.

Suponga que A es de $m \times n$ y B es de $n \times r$, de modo que existe el producto AB. Si particiona B en términos de sus vectores columnas, como $B = [\mathbf{b}_1 \vdots \mathbf{b}_2 \vdots \cdots \vdots \mathbf{b}_r]$, entonces

$$AB = A[\mathbf{b}_1 \vdots \mathbf{b}_2 \vdots \cdots \vdots \mathbf{b}_r] = [A\mathbf{b}_1 \vdots A\mathbf{b}_2 \vdots \cdots \vdots A\mathbf{b}_r]$$

Este resultado es una consecuencia inmediata de la definición de multiplicación de matrices. La forma a la derecha se llama ***representación matriz-columna*** del producto.

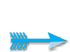
Ejemplo 3.9

Si

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix}$$

entonces

$$A\mathbf{b}_1 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 13 \\ 2 \end{bmatrix} \quad \text{y} \quad A\mathbf{b}_2 = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ -2 \end{bmatrix}$$



Por tanto, $AB = [A\mathbf{b}_1 \vdots A\mathbf{b}_2] = \begin{bmatrix} 13 & \vdots & 5 \\ 2 & \vdots & -2 \end{bmatrix}$. (Compruebe mediante multiplicación matricial ordinaria.)



Comentario Advierta que la representación matriz-columna de AB permite escribir cada columna de AB como una combinación lineal de las columnas de A con entradas de B como coeficientes. Por ejemplo,

$$\begin{bmatrix} 13 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

(Vea los ejercicios 23 y 26.)

Suponga que A es de $m \times n$ y B es de $n \times r$, de modo que existe el producto AB. Si particiona A en términos de sus vectores renglón como

$$A = \begin{bmatrix} \mathbf{A}_1 \\ \hdashline \mathbf{A}_2 \\ \hdashline \vdots \\ \hdashline \mathbf{A}_m \end{bmatrix}$$

entonces

$$AB = \begin{bmatrix} \mathbf{A}_1 \\ \hdashline \mathbf{A}_2 \\ \hdashline \vdots \\ \hdashline \mathbf{A}_m \end{bmatrix} B = \begin{bmatrix} \mathbf{A}_1 B \\ \hdashline \mathbf{A}_2 B \\ \hdashline \vdots \\ \hdashline \mathbf{A}_m B \end{bmatrix}$$

Una vez más, este resultado es consecuencia directa de la definición de multiplicación matricial. La forma a la derecha se conoce como ***representación renglón-matriz*** del producto.

Ejemplo 3.10

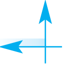
Con la representación renglón-matriz, calcule AB para las matrices del ejemplo 3.9.

Solución Calcule

$$\mathbf{A}_1B = [1 \quad 3 \quad 2]\begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix} = [13 \quad 5] \quad \text{y} \quad \mathbf{A}_2B = [0 \quad -1 \quad 1]\begin{bmatrix} 4 & -1 \\ 1 & 2 \\ 3 & 0 \end{bmatrix}$$

$$= [2 \quad -2]$$

Por tanto, $AB = \begin{bmatrix} \mathbf{A}_1B \\ \hline \mathbf{A}_2B \end{bmatrix} = \begin{bmatrix} 13 & 5 \\ \hline 2 & -2 \end{bmatrix}$, como antes



La definición del producto matricial AB usa la partición natural de A en renglones y B en columnas; esta forma bien puede llamarse ***representación renglón-columna*** del producto. También puede partir A en columnas y B en renglones; esta forma se llama ***representación columna-renglón*** del producto.

En este caso, se tiene

$$A = [\mathbf{a}_1 \vdots \mathbf{a}_2 \vdots \cdots \vdots \mathbf{a}_n] \quad \text{y} \quad B = \begin{bmatrix} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \\ \hline \vdots \\ \hline \mathbf{B}_n \end{bmatrix}$$

de modo que
$$AB = [\mathbf{a}_1 \vdots \mathbf{a}_2 \vdots \cdots \vdots \mathbf{a}_n]\begin{bmatrix} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \\ \hline \vdots \\ \hline \mathbf{B}_n \end{bmatrix} = \mathbf{a}_1\mathbf{B}_1 + \mathbf{a}_2\mathbf{B}_2 + \cdots + \mathbf{a}_n\mathbf{B}_n \tag{2}$$

Note que la suma recuerda una expansión de producto punto; la diferencia es que los términos individuales son matrices, no escalares. Asegúrese de que esto tiene sentido. Cada término $\mathbf{a}_i\mathbf{B}_i$ es el producto de una matriz de $m \times 1$ y una de $1 \times r$. Por tanto, cada $\mathbf{a}_i\mathbf{B}_i$ es una matriz de $m \times r$ del mismo tamaño que AB. Los productos $\mathbf{a}_i\mathbf{B}_i$ se llaman ***productos exteriores*** y (2) se llama ***expansión de producto exterior*** de AB.

Ejemplo 3.11

Calcule la expansión de producto exterior de AB para las matrices del ejemplo 3.9.

Solución Se tiene

$$A = [\mathbf{a}_1 \vdots \mathbf{a}_2 \vdots \mathbf{a}_3] = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \mathbf{B}_1 \\ \hline \mathbf{B}_2 \\ \hline \mathbf{B}_3 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ \hline 1 & 2 \\ \hline 3 & 0 \end{bmatrix}$$

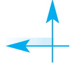
Los productos exteriores son

$$\mathbf{a}_1\mathbf{B}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}[4 \quad -1] = \begin{bmatrix} 4 & -1 \\ 0 & 0 \end{bmatrix}, \quad \mathbf{a}_2\mathbf{B}_2 = \begin{bmatrix} 3 \\ -1 \end{bmatrix}[1 \quad 2] = \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix}$$

y
$$\mathbf{a}_3\mathbf{B}_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}[3 \quad 0] = \begin{bmatrix} 6 & 0 \\ 3 & 0 \end{bmatrix}$$

(Observe que calcular cada producto exterior es exactamente igual que llenar una tabla de multiplicación.) Por tanto, la expansión de producto exterior de AB es

$$\mathbf{a}_1\mathbf{B}_1 + \mathbf{a}_2\mathbf{B}_2 + \mathbf{a}_3\mathbf{B}_3 = \begin{bmatrix} 4 & -1 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 6 \\ -1 & -2 \end{bmatrix} + \begin{bmatrix} 6 & 0 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 13 & 5 \\ 2 & -2 \end{bmatrix} = AB$$



La expansión de producto exterior se usará en los capítulos 5 y 7 cuando se estudie el Teorema Espectral y la descomposición de valor singular, respectivamente.

Cada una de las particiones siguientes es un caso especial del particionamiento en general. Se dice que una matriz A está particionada si se introdujeron rectas horizontales y verticales que subdividen A en submatrices llamadas bloques. El particionamiento permite que A se escriba como matriz cuyas entradas son sus bloques.

Por ejemplo,

$$A = \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 2 & -1 \\ 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 4 & 0 \\ \hline 0 & 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 7 & 2 \end{array}\right] \quad \text{y} \quad B = \left[\begin{array}{cc|cc|c} 4 & 3 & 1 & 2 & 1 \\ -1 & 2 & 2 & 1 & 1 \\ 1 & -5 & 3 & 3 & 1 \\ \hline 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 3 \end{array}\right]$$

son matrices particionadas. Tienen las estructuras de bloque

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix}$$

Si dos matrices tienen el mismo tamaño y se parten la misma forma, es claro que pueden sumarse y multiplicarse por escalares bloque por bloque. Menos obvio es el hecho de que, con el particionamiento adecuado, las matrices también pueden multiplicarse en forma de bloque. El siguiente ejemplo ilustra este proceso.

Ejemplo 3.12

Considere las matrices A y B anteriores. Si por el momento ignora el hecho de que sus entradas son matrices, entonces A parece ser una matriz de 2×2 y B una matriz de 2×3. En consecuencia, su producto debe ser una matriz de 2×3 dada por

$$\begin{aligned} AB &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \end{bmatrix} \\ &= \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} & A_{11}B_{13} + A_{12}B_{23} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} & A_{21}B_{13} + A_{22}B_{23} \end{bmatrix} \end{aligned}$$

Pero todos los productos de este cálculo en realidad son productos *matriciales*, así que debe asegurarse de que todos estén definidos. Una comprobación rápida revela que de hecho este es el caso, pues los números de *columnas* en los bloques de A (3 y 2) coinciden con los números de *renglones* en los bloques de B. Se dice que las matrices A y B están ***particionadas de conformidad para multiplicación por bloques***.

Al realizar los cálculos indicados se obtiene el producto AB en forma particionada:

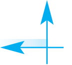
$$A_{11}B_{11} + A_{12}B_{21} = I_3B_{11} + A_{12}I_2 = B_{11} + A_{12} = \begin{bmatrix} 4 & 3 \\ -1 & 2 \\ 1 & -5 \end{bmatrix} + \begin{bmatrix} 2 & -1 \\ 1 & 3 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 0 & 5 \\ 5 & -5 \end{bmatrix}$$

(Cuando algunos bloques son matrices cero o matrices identidad, como es el caso aquí, dichos cálculos pueden realizarse muy rápidamente.) Los cálculos para los otros cinco bloques de AB son similares. Compruebe que el resultado es

$$\left[\begin{array}{cc|cc|c} 6 & 2 & 1 & 2 & 2 \\ 0 & 5 & 2 & 1 & 12 \\ 5 & -5 & 3 & 3 & 9 \\ \hline 1 & 7 & 0 & 0 & 23 \\ 7 & 2 & 0 & 0 & 20 \end{array}\right]$$



(Observe que el bloque en la esquina superior izquierda es el resultado de los cálculos anteriores.) Compruebe que obtiene la misma respuesta al multiplicar A por B en la forma usual.



Potencias de matrices

Cuando A y B son dos matrices de $n \times n$, su producto AB también será una matriz de $n \times n$. Un caso especial ocurre cuando $A = B$. Tiene sentido definir $A^2 = AA$ y, en general, definir A^k como

$$A^k = \underbrace{AA \cdots A}_{k \text{ factores}}$$

si k es un entero positivo. Por ende, $A^1 = A$ y es conveniente definir $A^0 = I_n$.

Antes de hacer demasiadas suposiciones, debe preguntarse en qué medida las potencias de matrices se comportan como potencias de números reales. Las siguientes propiedades se deducen inmediatamente de las definiciones recién dadas y son los análogos matriciales de las correspondientes propiedades de las potencias de los números reales.

Si A es una matriz cuadrada y r y s son enteros no negativos, entonces

1. $A^r A^s = A^{r+s}$
2. $(A^r)^s = A^{rs}$

En la sección 3.3 se extenderá la definición y las propiedades para incluir potencias de enteros negativos.

Ejemplo 3.13

(a) Si $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, entonces

$$A^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, \quad A^3 = A^2A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 4 \\ 4 & 4 \end{bmatrix}$$

y, en general,

$$A^n = \begin{bmatrix} 2^{n-1} & 2^{n-1} \\ 2^{n-1} & 2^{n-1} \end{bmatrix} \quad \text{para todo } n \geq 1$$

El enunciado anterior puede probarse mediante inducción matemática, pues es una colección *infinita* de enunciados, uno para cada número natural n. (El Apéndice B ofrece

un breve repaso de inducción matemática.) El paso básico es probar que la fórmula se cumple para $n = 1$. En este caso,

$$A^1 = \begin{bmatrix} 2^{1-1} & 2^{1-1} \\ 2^{1-1} & 2^{1-1} \end{bmatrix} = \begin{bmatrix} 2^0 & 2^0 \\ 2^0 & 2^0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = A$$

como se requiere.

La hipótesis de inducción es suponer que

$$A^k = \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{bmatrix}$$

para algún entero $k \geq 1$. El paso de inducción es probar que la fórmula se cumple para $n = k + 1$. Con la definición de potencias de matrices y la hipótesis de inducción, se calcula

$$\begin{aligned} A^{k+1} = A^kA &= \begin{bmatrix} 2^{k-1} & 2^{k-1} \\ 2^{k-1} & 2^{k-1} \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \\ 2^{k-1} + 2^{k-1} & 2^{k-1} + 2^{k-1} \end{bmatrix} \\ &= \begin{bmatrix} 2^k & 2^k \\ 2^k & 2^k \end{bmatrix} \\ &= \begin{bmatrix} 2^{(k+1)-1} & 2^{(k+1)-1} \\ 2^{(k+1)-1} & 2^{(k+1)-1} \end{bmatrix} \end{aligned}$$

Por tanto, la fórmula se cumple para toda $n \geq 1$ por el principio de inducción matemática.

(b) Si $B = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, entonces $B^2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$. Al continuar, se encuentra

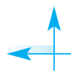
$$B^3 = B^2B = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

y

$$B^4 = B^3B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Por tanto, $B^5 = B$, y la sucesión de potencias de B se repite en un ciclo de cuatro:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \ldots$$



La transpuesta de una matriz

Hasta el momento, todas las operaciones matriciales definidas son análogas a las operaciones sobre números reales, aunque no siempre pueden comportarse en la misma forma. La siguiente operación no tiene tal análogo.

Definición La ***transpuesta*** de una matriz A de $m \times n$ es la matriz A^T de $n \times m$ que se obtiene al intercambiar las renglones y columnas de A. Esto es: la i-ésima columna de A^T es el i-ésimo renglón de A para toda i.

Ejemplo 3.14

Sea

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 5 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{y} \quad C = \begin{bmatrix} 5 & -1 & 2 \end{bmatrix}$$

Entonces sus transpuestas son

$$A^T = \begin{bmatrix} 1 & 5 \\ 3 & 0 \\ 2 & 1 \end{bmatrix}, \quad B^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \quad \text{y} \quad C^T = \begin{bmatrix} 5 \\ -1 \\ 2 \end{bmatrix}$$

En ocasiones, la transpuesta se usa para dar una definición alternativa del producto punto de dos vectores en términos de multiplicación de matrices. Si

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

entonces

$$\begin{aligned} \mathbf{u} \cdot \mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n \\ &= \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \\ &= \mathbf{u}^T\mathbf{v} \end{aligned}$$

Una definición alternativa útil de la transpuesta se da en forma de componentes:

$$(A^T)_{ij} = A_{ji} \quad \text{para toda } i \text{ y } j$$

En palabras: las entradas en el renglón i y la columna j de A^T son iguales que las entradas en el renglón j y la columna i de A.

La transpuesta también se usa para definir un tipo muy importante de matriz cuadrada: una matriz simétrica.

Definición Una matriz cuadrada A es ***simétrica*** si $A^T = A$; esto es: si A es igual a su propia transpuesta.

Ejemplo 3.15

Sea

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 5 & 0 \\ 2 & 0 & 4 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$$

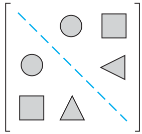
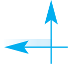


Figura 3.1
Una matriz simétrica

Entonces A es simétrica, pues $A^T = A$; pero B no es simétrica, pues $B^T = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix} \neq B$.



Una matriz simétrica tiene la propiedad de que es su propia "imagen especular" a través de su diagonal principal. La figura 3.1 ilustra esta propiedad para una matriz de 3×3. Las formas correspondientes representan entradas iguales; las entradas de la diagonal (las que están sobre la recta rayada) son arbitrarias.

Una definición por componentes de una matriz simétrica también es útil. Es simplemente la descripción algebraica de la propiedad de "reflexión".

Una matriz cuadrada A es simétrica si y sólo si $A_{ij} = A_{ji}$ para todo i y j.

Ejercicios 3.1

Sean

$$A = \begin{bmatrix} 3 & 0 \\ -1 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -2 & 1 \\ 0 & 2 & 3 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & -3 \\ -2 & 1 \end{bmatrix}, \quad E = \begin{bmatrix} 4 & 2 \end{bmatrix}, \quad F = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

En los ejercicios 1-16, calcule las matrices indicadas (si es posible).

1. $A + 2D$

2. $2D - 5A$

3. $B - C$

4. $B - C^T$

5. AB

6. B^2

7. $D + BC$

8. B^TB

9. $E(AF)$

10. $F(AF)$

11. FE

12. EF

13. $B^TC^T - (CB)^T$

14. $DA - AD$

15. A^3

16. $(I_2 - A)^2$

17. Proporcione un ejemplo de una matriz A de 2×2 distinta de cero tal que $A^2 = 0$.

18. Sea $A = \begin{bmatrix} 2 & 1 \\ 6 & 3 \end{bmatrix}$. Encuentre matrices de 2×2 B y C tales que $AB = AC$, pero $B \neq C$.

19. Una fábrica elabora tres productos (chismes, cachivaches y chunches) y los embarca a dos almacenes para su cuidado. El número de unidades de cada producto embarcado a cada almacén está dado por la matriz

$$A = \begin{bmatrix} 200 & 75 \\ 150 & 100 \\ 100 & 125 \end{bmatrix}$$

(donde a_{ij} es el número de unidades de producto i enviadas al almacén j y los productos se guardan en orden alfabético). El costo de embarcar una unidad de cada producto por camión es \$1.50 por chisme, \$1.00 por cachivache y \$2.00 por chunche. Los correspondientes costos unitarios para embarcar por tren son \$1.75, \$1.50 y \$1.00. Organice estos costos en una matriz B y luego use la multiplicación de matrices para mostrar cómo la fábrica puede comparar el costo de embarcar sus productos a cada uno de los dos almacenes por camión y por tren.

20. Con referencia al ejercicio 19, suponga que el costo unitario de distribuir los productos a tiendas es el mismo para cada producto, pero varía por almacén debido a las distancias involucradas. Cuesta \$0.75 distribuir una unidad desde el almacén 1 y \$1.00 distribuir una unidad desde el almacén 2. Organice dichos costos en una matriz C y luego use la multiplicación de matrices para calcular el costo total de distribuir cada producto.

En los ejercicios 21-22, escriba el sistema dado de ecuaciones lineales como una ecuación matricial de la forma $A\mathbf{x} = \mathbf{b}$.

21. $$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 0 \\ 2x_1 + x_2 - 5x_3 &= 4 \end{aligned}$$

22. $$\begin{aligned} -x_1 \phantom{{}-x_2} + 2x_3 &= 1 \\ x_1 - x_2 \phantom{{}+2x_3} &= -2 \\ x_2 + x_3 &= -1 \end{aligned}$$

En los ejercicios 23-28, sean

$$A = \begin{bmatrix} 1 & 0 & -2 \\ -3 & 1 & 1 \\ 2 & 0 & -1 \end{bmatrix}$$

y

$$B = \begin{bmatrix} 2 & 3 & 0 \\ 1 & -1 & 1 \\ -1 & 6 & 4 \end{bmatrix}$$

23. Use la representación matriz-columna del producto para escribir cada columna de AB como una combinación lineal de las columnas de A.

24. Use la representación renglón-matriz del producto para escribir cada renglón de AB como una combinación lineal de los renglones de B.

25. Calcule el desarrollo del producto exterior de AB.

26. Use la representación matriz-columna del producto para escribir cada columna de BA como una combinación lineal de las columnas de B.

27. Use la representación renglón-matriz del producto para escribir cada renglón de BA como una combinación lineal de los renglones de A.

28. Calcule el desarrollo del producto exterior de BA.

En los ejercicios 29 y 30, suponga que el producto AB tiene sentido.

29. Demuestre que, si las columnas de B son linealmente dependientes, entonces también lo son las columnas de AB.

30. Demuestre que, si los renglones de A son linealmente dependientes, entonces también lo son los renglones de AB.

En los ejercicios 31-34, calcule AB mediante multiplicación por bloques y use la partición indicada.

31. $A = \left[\begin{array}{cc|cc} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & 3 \end{array}\right]$, $B = \left[\begin{array}{cc|c} 2 & 3 & 0 \\ -1 & 1 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 1 \end{array}\right]$

32. $A = \left[\begin{array}{cc|cc} 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{array}\right]$, $B = \left[\begin{array}{c|cc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \hline 1 & 5 & 4 \\ -2 & 3 & 2 \end{array}\right]$

33. $A = \left[\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \\ \hline 1 & 0 & -1 & 1 \\ 0 & 1 & 1 & -1 \end{array}\right]$, $B = \left[\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{array}\right]$

34. $A = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \\ \hline 0 & 0 & 0 & 4 \end{array}\right]$, $B = \left[\begin{array}{ccc|c} 1 & 2 & 3 & 1 \\ 0 & 1 & 4 & 1 \\ 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & -1 \end{array}\right]$

35. Sea $A = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$.

(a) Calcule $A^2, A^3, \ldots, A^7$.

(b) ¿Qué es A^{2001}? ¿Por qué?

36. Sea $B = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. Encuentre y justifique B^{2011}.

37. Sea $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$. Encuentre una fórmula para A^n $(n \geq 1)$ y verifique su fórmula usando inducción matemática.

38. Sea $A = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$.

(a) Demuestre que $A^2 = \begin{bmatrix} \cos 2\theta & -\operatorname{sen} 2\theta \\ \operatorname{sen} 2\theta & \cos 2\theta \end{bmatrix}$.

(b) Demuestre, mediante inducción matemática, que

$$A^n = \begin{bmatrix} \cos n\theta & -\operatorname{sen} n\theta \\ \operatorname{sen} n\theta & \cos n\theta \end{bmatrix} \text{ para } n \geq 1$$

39. En cada uno de los siguientes, encuentre la matriz $A = [a_{ij}]$ de 4×4 que satisfaga la condición dada:

(a) $a_{ij} = (-1)^{i+j}$ **(b)** $a_{ij} = j - i$

(c) $a_{ij} = (i - 1)^j$ **(d)** $a_{ij} = \operatorname{sen}\left(\dfrac{(i + j - 1)\pi}{4}\right)$

40. En cada uno de los siguientes, encuentre la matriz $A = [a_{ij}]$ de 6×6 que satisfaga la condición dada:

(a) $a_{ij} = \begin{cases} i + j & \text{si } i \leq j \\ 0 & \text{si } i > j \end{cases}$ **(b)** $a_{ij} = \begin{cases} 1 & \text{si } |i - j| \leq 1 \\ 0 & \text{si } |i - j| > 1 \end{cases}$

(c) $a_{ij} = \begin{cases} 1 & \text{si } 6 \leq i + j \leq 8 \\ 0 & \text{de otro modo} \end{cases}$

41. Demuestre el Teorema 3.1(a).

3.2 Álgebra matricial

En ciertos aspectos, la aritmética de matrices generaliza la de los vectores. No se esperan sorpresas con respecto a la suma y la multiplicación por un escalar, y de hecho no las hay. Esto permitirá extender a las matrices varios conceptos que ya son familiares a partir del trabajo con vectores. En particular, combinaciones lineales, conjuntos generadores e independencia lineal se trasladan a las matrices sin dificultad.

Sin embargo, las matrices tienen otras operaciones, como la multiplicación matricial, que no poseen los vectores. No debe esperar que la multiplicación matricial se comporte como la multiplicación de números reales, a menos que pueda probar que lo hace; de hecho, no lo hace. En esta sección se resumen y demuestran algunas de las propiedades principales de las operaciones matriciales y se comienza por desarrollar un álgebra de matrices.

Propiedades de la suma y la multiplicación por un escalar

Todas las propiedades algebraicas de la suma y la multiplicación por un escalar para vectores (Teorema 1.1) se trasladan a las matrices. En el siguiente teorema se resumen dichas propiedades.

Teorema 3.2 **Propiedades algebraicas de la suma y la multiplicación escalar matriciales**

Sean A, B y C matrices del mismo tamaño, y sean c y d escalares. Entonces

a. $A + B = B + A$ — Conmutatividad
b. $(A + B) + C = A + (B + C)$ — Asociatividad
c. $A + O = A$
d. $A + (-A) = O$
e. $c(A + B) = cA + cB$ — Distributividad
f. $(c + d)A = cA + dA$ — Distributividad
g. $c(dA) = (cd)A$
h. $1A = A$

Las demostraciones de dichas propiedades son análogos directos de las correspondientes pruebas de las propiedades vectoriales y se dejan como ejercicios. Del mismo modo, los comentarios que siguen al Teorema 1.1 son igualmente válidos aquí y usted no debe tener dificultad para usar dichas propiedades en la realización de manipulaciones algebraicas con matrices. (Revise el ejemplo 1.5 y vea los ejercicios 17 y 18 al final de esta sección.)

La propiedad de asociatividad permite sin ambigüedades combinar la multiplicación por un escalar y la suma sin paréntesis. Si A, B y C son matrices del mismo tamaño, entonces

$$(2A + 3B) - C = 2A + (3B - C)$$

y por tanto simplemente puede escribirse $2A + 3B - C$. En general, entonces, si A_1, $A_2, \dots, A_k$ son matrices del mismo tamaño y $c_1, c_2, \dots, c_k$ son escalares, puede formarse la ***combinación lineal***

$$c_1A_1 + c_2A_2 + \cdots + c_kA_k$$

A $c_1, c_2, \dots, c_k$ se les denomina como los ***coeficientes*** de la combinación lineal. Ahora puede plantear y responder preguntas acerca de las combinaciones lineales de matrices.

Ejemplo 3.16

Sean $A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $A_3 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$.

(a) ¿$B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ es una combinación lineal de A_1, A_2 y A_3?

(b) ¿$C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ es una combinación lineal de A_1, A_2 y A_3?

Solución

(a) Se quieren encontrar escalares c_1, c_2 y c_3 tales que $c_1A_1 + c_2A_2 + c_3A_3 = B$. Por tanto,

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$$

El lado izquierdo de esta ecuación puede reescribirse como

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix}$$

Al comparar entradas y usar la definición de igualdad matricial, se tienen cuatro ecuaciones lineales

$$\begin{aligned} c_2 + c_3 &= 1 \\ c_1 \qquad + c_3 &= 4 \\ -c_1 \qquad + c_3 &= 2 \\ c_2 + c_3 &= 1 \end{aligned}$$

La eliminación de Gauss-Jordan produce fácilmente

$$\left[\begin{array}{rrr|r} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{rrr|r} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$



(¡compruébelo!), de modo que $c_1 = 1$, $c_2 = -2$, y $c_3 = 3$. En consecuencia, $A_1 - 2A_2 + 3A_3 = B$, que puede comprobarse fácilmente.

(b) Esta vez se quiere resolver

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

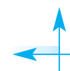
Al proceder como en el inciso (a), se obtiene el sistema lineal

$$\begin{aligned} c_2 + c_3 &= 1 \\ c_1 \qquad + c_3 &= 2 \\ -c_1 \qquad + c_3 &= 3 \\ c_2 + c_3 &= 4 \end{aligned}$$

La reducción por renglones produce

$$\left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 4 \end{array}\right] \xrightarrow{R_4 - R_1} \left[\begin{array}{ccc|c} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 2 \\ -1 & 0 & 1 & 3 \\ 0 & 0 & 0 & 3 \end{array}\right]$$

No es necesario avanzar más: el último renglón implica que no hay solución. Por tanto, en este caso, C no es una combinación lineal de A_1, A_2 y A_3.



Comentario Observe que las columnas de la matriz aumentada contienen las entradas de las matrices que se proporcionan. Si se leen las entradas de cada matriz de izquierda a derecha y de arriba abajo, se obtiene el orden en el que aparecen en las columnas de la matriz aumentada. Por ejemplo, A_1 se lee como "0,1,−1,0", que corresponde a la primera columna de la matriz aumentada. Es como si simplemente se "estiraran" las matrices dadas en vectores columna. Por tanto, habría terminado exactamente con el mismo sistema de ecuaciones lineales que en el inciso (a) si se hubiera preguntado:

$$\text{Es } \begin{bmatrix} 1 \\ 4 \\ 2 \\ 1 \end{bmatrix} \text{ una combinación lineal de } \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \text{ y } \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}?$$

A partir de ahora se encontrarán con frecuencia tales paralelismos. En el capítulo 6 se les explorará con más detalle.

Se puede definir el ***espacio generado*** por un conjunto de matrices como el conjunto de todas las combinaciones lineales de las matrices.

Ejemplo 3.17

Describa el espacio generado por las matrices A_1, A_2 y A_3 en el ejemplo 3.16.

Solución Una forma de hacer esto es simplemente escribir una combinación lineal general de A_1, A_2 y A_3. Por tanto,

$$\begin{aligned} c_1A_1 + c_2A_2 + c_3A_3 &= c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} \end{aligned}$$

(que es análogo a la representación paramétrica de un plano). Pero suponga que se quiere saber cuándo la matriz $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ está en gen(A_1, A_2, A_3). A partir de la representación anterior, se sabe que lo está cuando

$$\begin{bmatrix} c_2 + c_3 & c_1 + c_3 \\ -c_1 + c_3 & c_2 + c_3 \end{bmatrix} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$$

para alguna elección de escalares c_1, c_2, c_3. Esto da origen a un sistema de ecuaciones lineales cuyo lado izquierdo es exactamente igual que el del ejemplo 3.16, pero cuyo lado derecho es general. La matriz aumentada de este sistema es

$$\left[\begin{array}{rrr|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right]$$

y la reducción por renglones produce

$$\left[\begin{array}{rrr|c} 0 & 1 & 1 & w \\ 1 & 0 & 1 & x \\ -1 & 0 & 1 & y \\ 0 & 1 & 1 & z \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{2}x - \frac{1}{2}y \\ 0 & 1 & 0 & -\frac{1}{2}x - \frac{1}{2}y + w \\ 0 & 0 & 1 & \frac{1}{2}x + \frac{1}{2}y \\ 0 & 0 & 0 & w - z \end{array}\right]$$



(Compruebe esto cuidadosamente.) La única restricción proviene del último renglón, donde claramente se debe tener $w - z = 0$ para tener una solución. Por tanto, el gen de A_1, A_2 y A_3 consiste de todas las matrices $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ para las cuales $w = z$. Esto es es: $\text{gen}(A_1, A_2, A_3) = \left\{\begin{bmatrix} w & x \\ y & w \end{bmatrix}\right\}$.



Nota Si se hubiese sabido esto *antes* de intentar el ejemplo 3.16, habría visto inmediatamente que $B = \begin{bmatrix} 1 & 4 \\ 2 & 1 \end{bmatrix}$ es una combinación lineal de A_1, A_2 y A_3, pues tiene la forma necesaria (tome $w = 1, x = 4$ y $y = 2$), pero $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ no puede ser una combinación lineal de A_1, A_2 y A_3, ya que no tiene la forma adecuada ($1 \neq 4$).

La independencia lineal también tiene sentido para matrices. Se dice que las matrices $A_1, A_2, \ldots, A_k$ del mismo tamaño son ***linealmente independientes*** si la única solución de la ecuación

$$c_1A_1 + c_2A_2 + \cdots + c_kA_k = O \tag{1}$$

es la trivial: $c_1 = c_2 = \cdots = c_k = 0$. Si no hay coeficientes triviales que satisfagan (1), entonces $A_1, A_2, \ldots, A_k$ se llama ***linealmente dependiente***.

Ejemplo 3.18

Determine si las matrices A_1, A_2 y A_3 en el ejemplo 3.16 son linealmente independientes.

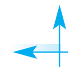
Solución Se quiere resolver la ecuación $c_1A_1 + c_2A_2 + c_3A_3 = O$. Al escribir las matrices, se tiene

$$c_1\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} + c_2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + c_3\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Esta vez se tiene un sistema lineal *homogéneo* cuyo lado izquierdo es el mismo que en los ejemplos 3.16 y 3.17. (¿Comienza a vislumbrar un patrón?) La matriz aumentada se reduce por renglones para obtener

$$\left[\begin{array}{rrr|c} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Por tanto, $c_1 = c_2 = c_3 = 0$, y se concluye que las matrices A_1, A_2 y A_3 son linealmente independientes.



Propiedades de la multiplicación matricial

Siempre que encuentre una nueva operación, como la multiplicación matricial, debe tener cuidado de no suponer demasiado acerca de ella. Sería bueno si la multiplicación matricial se comportara como la multiplicación de números reales. Aunque en muchos aspectos lo hace, existen algunas diferencias significativas.

Ejemplo 3.19

Considere las matrices

$$A = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Al multiplicar se obtiene

$$AB = \begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 4 \\ -3 & -2 \end{bmatrix} \quad \text{y} \quad BA = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 4 \\ -1 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$$

Por tanto, $AB \neq BA$. Así que, en contraste con la multiplicación de números reales, la multiplicación de matrices *no es conmutativa*: ¡el orden de los factores en un producto sí importa!

Es fácil comprobar que $A^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (¡hágalo!). De este modo, para matrices, la



ecuación $A^2 = O$ no implica que $A = O$ (a diferencia de la situación para números reales, donde la ecuación $x^2 = 0$ sólo tiene $x = 0$ como solución).



Por muy tristes que parezcan las cosas después del último ejemplo, la situación realmente no es tan mala: sólo necesita acostumbrarse a trabajar con matrices y recordar constantemente que no son números. El siguiente teorema resume las principales propiedades de la multiplicación de matrices.

Teorema 3.3 **Propiedades de la multiplicación de matrices**

Sean A, B y C matrices (cuyos tamaños son tales que pueden realizarse las operaciones indicadas) y sea k un escalar. Entonces

a. $A(BC) = (AB)C$ — Asociativa
b. $A(B + C) = AB + AC$ — Distributiva izquierda
c. $(A + B)C = AC + BC$ — Distributiva derecha
d. $k(AB) = (kA)B = A(kB)$
e. $I_mA = A = AI_n$ si A es $m \times n$ — Identidad multiplicativa

Demostración Se demuestran (b) y la mitad de (e). La demostración de la propiedad (a) se difiere hasta la sección 3.6. Las propiedades restantes se consideran en los ejercicios.

(b) Para demostrar $A(B + C) = AB + AC$, los renglones de A se denotan como $\mathbf{A}_i$ y las columnas de B y C como $\mathbf{b}_j$ y $\mathbf{c}_j$. Entonces la columna j-ésima de $B + C$ es $\mathbf{b}_j + \mathbf{c}_j$ (ya que la suma está definida por componentes) y por tanto

$$\begin{aligned}[A(B + C)]_{ij} &= \mathbf{A}_i \cdot (\mathbf{b}_j + \mathbf{c}_j)\\ &= \mathbf{A}_i \cdot \mathbf{b}_j + \mathbf{A}_i \cdot \mathbf{c}_j\\ &= (AB)_{ij} + (AC)_{ij}\\ &= (AB + AC)_{ij}\end{aligned}$$

Dado que esto es cierto para todo i y j, se debe tener $A(B + C) = AB + AC$.
(e) Para demostrar $AI_n = A$, note que la matriz identidad I_n puede partirse en columnas como

$$I_n = [\mathbf{e}_1 \vdots \mathbf{e}_2 \vdots \cdots \vdots \mathbf{e}_n]$$

donde $\mathbf{e}_i$ es un vector unitario estándar. Por tanto,

$$\begin{aligned}AI_n &= [A\mathbf{e}_1 \vdots A\mathbf{e}_2 \vdots \cdots \vdots A\mathbf{e}_n]\\ &= [\mathbf{a}_1 \vdots \mathbf{a}_2 \vdots \cdots \vdots \mathbf{a}_n]\\ &= A\end{aligned}$$

por el Teorema 3.1(b).

Puede usar estas propiedades para explorar aún más cuán estrechamente la multiplicación de matrices recuerda la multiplicación de números reales.

Ejemplo 3.20

Si A y B son matrices cuadradas del mismo tamaño, ¿$(A + B)^2 = A^2 + 2AB + B^2$?

Solución Con las propiedades de la multiplicación de matrices, se calcula

$$\begin{aligned}(A + B)^2 &= (A + B)(A + B)\\ &= (A + B)A + (A + B)B && \text{por distributividad izquierda}\\ &= A^2 + BA + AB + B^2 && \text{por distributividad derecha}\end{aligned}$$

Por tanto, $(A + B)^2 = A^2 + 2AB + B^2$ si y sólo si $A^2 + BA + AB + B^2 = A^2 + 2AB + B^2$. Al restar A^2 y B^2 de ambos lados se obtiene $BA + AB = 2AB$. Al restar AB de ambos lados se obtiene $BA = AB$. Por tanto, $(A + B)^2 = A^2 + 2AB + B^2$ si y sólo si A y B se conmutan. (¿Puede dar un ejemplo de tal par de matrices? ¿Puede encontrar dos matrices que no satisfagan esta propiedad?)

Propiedades de la transpuesta

Teorema 3.4

Propiedades de la transpuesta

Sean A y B matrices (cuyos tamaños son tales que pueden realizarse las operaciones indicadas) y sea k un escalar. Entonces

a. $(A^T)^T = A$
b. $(A + B)^T = A^T + B^T$
c. $(kA)^T = k(A^T)$
d. $(AB)^T = B^TA^T$
e. $(A^r)^T = (A^T)^r$ para todos los enteros r no negativos

Demostración Las propiedades (a)-(c) son intuitivamente claras y directas de demostrar (vea el ejercicio 30). Demostrar la propiedad (e) es un buen ejercicio de inducción matemática (vea el ejercicio 31). Se demostrará (d), pues no es lo que se hubiera esperado. [¿Usted sospecharía que $(AB)^T = A^TB^T$ puede ser verdadero?]

Primero, si A es de $m \times n$ y B es de $n \times r$, entonces B^T es de $r \times n$ y A^T es de $n \times m$. Por tanto, el producto B^TA^T está definido y es de $r \times m$. Dado que AB es de $m \times r$, $(AB)^T$ es de $r \times m$, y por tanto $(AB)^T$ y B^TA^T tienen el mismo tamaño. Ahora debe demostrar que sus correspondientes entradas son iguales.

Denote el i-ésimo renglón de una matriz X mediante $\text{renglón}_i(X)$ y su j-ésima columna mediante $\text{col}_j(X)$. Al usar estas convenciones, se ve que

$$\begin{aligned}[(AB)^T]_{ij} &= (AB)_{ji} \\ &= \text{renglón}_j(A) \cdot \text{col}_i(B) \\ &= \text{col}_j(A^T) \cdot \text{renglón}_i(B^T) \\ &= \text{renglón}_i(B^T) \cdot \text{col}_j(A^T) = [B^TA^T]_{ij}\end{aligned}$$

(Note que se usó la definición de multiplicación de matrices, la definición de la transpuesta y el hecho de que el producto punto es conmutativo.) Dado que i y j son arbitrarios, este resultado implica que $(AB)^T = B^TA^T$.

Comentario Las propiedades (b) y (d) del Teorema 3.4 pueden generalizarse a sumas y productos de un número finito de matrices:

$$(A_1 + A_2 + \cdots + A_k)^T = A_1^T + A_2^T + \cdots + A_k^T \quad \text{y} \quad (A_1A_2\cdots A_k)^T = A_k^T \cdots A_2^TA_1^T$$

si supone que los tamaños de las matrices son tales que todas las operaciones pueden realizarse. En los ejercicios 32 y 33 se le pide demostrar estos hechos mediante inducción matemática.

Ejemplo 3.21

Sean

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix}$$

Entonces $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$, de modo que $A + A^T = \begin{bmatrix} 2 & 5 \\ 5 & 8 \end{bmatrix}$, una matriz simétrica.

Se tiene

$$B^T = \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix}$$

de modo que

$$BB^T = \begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 5 \\ 5 & 14 \end{bmatrix}$$

y

$$B^TB = \begin{bmatrix} 4 & 2 \\ -1 & 3 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 4 & -1 & 0 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 20 & 2 & 2 \\ 2 & 10 & 3 \\ 2 & 3 & 1 \end{bmatrix}$$

En consecuencia, tanto BB^T como B^TB son simétricas, ¡aun cuando B ni siquiera es cuadrada! (Compruebe que AA^T y A^TA también son simétricas.)

El siguiente teorema dice que los resultados del ejemplo 3.21 son verdaderos en general.

Teorema 3.5 a. Si A es una matriz cuadrada, entonces $A + A^T$ es una matriz simétrica.
b. Para cualquier matriz A, AA^T y A^TA son matrices simétricas.

Demostración Se demuestra (a) y se deja la demostración de (b) como ejercicio 34. Simplemente demuestre que

$$(A + A^T)^T = A^T + (A^T)^T = A^T + A = A + A^T$$

(con las propiedades de la transpuesta y la conmutatividad de la suma de matrices). Por tanto, $A + A^T$ es igual a su propia transpuesta y en consecuencia, por definición, es simétrica.

Ejercicios 3.2

En los ejercicios 1-4, resuelva la ecuación para X, dado que

$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ *y* $B = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$.

1. $X - 2A + 3B = O$

2. $3X = A - 2B$

3. $2(A + 2B) = 3X$

4. $2(A - B + 2X) = 3(X - B)$

En los ejercicios 5-8, escriba B como una combinación lineal de las otras matrices, si es posible.

5. $B = \begin{bmatrix} 2 & 5 \\ 0 & 3 \end{bmatrix}$, $A_1 = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$

6. $B = \begin{bmatrix} 2 & -1 \\ -3 & 2 \end{bmatrix}$, $A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

7. $B = \begin{bmatrix} 3 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$, $A_1 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$, $A_2 = \begin{bmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

8. $B = \begin{bmatrix} 6 & -2 & 5 \\ -2 & 8 & 6 \\ 5 & 6 & 6 \end{bmatrix}$, $A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $A_2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, $A_4 = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix}$

En los ejercicios 9-12, encuentre la forma general del generador de las matrices indicadas, como en el ejemplo 3.17.

9. gen(A_1, A_2) en el ejercicio 5

10. gen(A_1, A_2, A_3) en el ejercicio 6

11. gen(A_1, A_2, A_3) en el ejercicio 7

12. gen(A_1, A_2, A_3, A_4) en el ejercicio 8

En los ejercicios 13-16, determine si las matrices dadas son linealmente independientes

13. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 4 & 3 \\ 2 & 1 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 0 & 1 \\ 5 & 2 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} -2 & -1 \\ 0 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} -1 & -3 \\ 1 & 9 \\ 4 & 5 \end{bmatrix}$

16. $\begin{bmatrix} 1 & -1 & 2 \\ 0 & 2 & 6 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 & 3 \\ 0 & 4 & 9 \\ 0 & 0 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 1 \\ 0 & 3 & 5 \\ 0 & 0 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 1 & -1 \\ 0 & 0 & -4 \\ 0 & 0 & 0 \end{bmatrix}$

17. Demuestre el Teorema 3.2(a)-(d).

18. Demuestre el Teorema 3.2(e)-(h).

19. Demuestre el Teorema 3.3(c).

20. Demuestre el Teorema 3.3(d).

21. Demuestre la mitad del Teorema 3.3(e) que no se demostró en el texto.

22. Demuestre que, para matrices cuadradas A y B, $AB = BA$ si y sólo si $(A - B)(A + B) = A^2 - B^2$.

En los ejercicios 23-25, si $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, *encuentre las condiciones de a, b, c y d para las que* $AB = BA$.

23. $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ **24.** $A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ **25.** $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

26. Encuentre las condiciones de a, b, c y d para los que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ conmuta tanto con $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ como con $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$.

27. Encuentre las condiciones sobre a, b, c y d tales que $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ conmuta con toda matriz de 2×2.

28. Demuestre que, si AB y BA están ambas definidas, entonces AB y BA son ambas matrices cuadradas.

Una matriz cuadrada se llama ***triangular superior*** *si todas las entradas abajo de la diagonal principal son cero. Por tanto, la forma de una matriz triangular superior es*

$$\begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & * & * \\ 0 & 0 & \cdots & 0 & * \end{bmatrix}$$

donde las entradas marcadas $*$ *son arbitrarias. Una definición más formal de tal matriz* $A = [a_{ij}]$ *es que* $a_{ij} = 0$ *si* $i > j$.

29. Demuestre que el producto de dos matrices triangulares superiores $n \times n$ es triangular superior.

30. Demuestre el Teorema 3.4(a)-(c).

31. Demuestre el Teorema 3.4(e).

32. Mediante inducción, demuestre que para toda $n \geq 1$, $(A_1 + A_2 + \cdots + A_n)^T = A_1^T + A_2^T + \cdots + A_n^T$.

33. Mediante inducción, demuestre que, para todo $n \geq 1$, $(A_1 A_2 \cdots A_n)^T = A_n^T \cdots A_2^T A_1^T$.

34. Demuestre el Teorema 3.5(b).

35. **(a)** Demuestre que si A y B son matrices simétricas de $n \times n$, entonces también lo es $A + B$.
(b) Demuestre que si A es una matriz simétrica de $n \times n$, entonces también lo es kA para cualquier escalar k.

36. **(a)** Proporcione un ejemplo para demostrar que si A y B son matrices simétricas de $n \times n$, entonces AB no necesita ser simétrica.
(b) Demuestre que, si A y B son matrices simétricas de $n \times n$, entonces AB es simétrica si y sólo si $AB = BA$.

Una matriz cuadrada se llama ***antisimétrica*** *si* $A^T = -A$.

37. ¿Cuáles de las siguientes matrices son antisimétricas?

(a) $\begin{bmatrix} 1 & 2 \\ -2 & 3 \end{bmatrix}$ **(b)** $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

(c) $\begin{bmatrix} 0 & 3 & -1 \\ -3 & 0 & 2 \\ 1 & -2 & 0 \end{bmatrix}$ **(d)** $\begin{bmatrix} 0 & 1 & 2 \\ -1 & 0 & 5 \\ 2 & 5 & 0 \end{bmatrix}$

38. Proporcione una definición en componentes de una matriz antisimétrica.

39. Demuestre que la diagonal principal de una matriz antisimétrica debe consistir completamente de ceros.

40. Demuestre que, si A y B son matrices de $n \times n$ antisimétricas, entonces también lo es $A + B$.

41. Si A y B son matrices de $n \times n$ antisimétricas, ¿en qué condiciones AB es antisimétrica?

42. Demuestre que, si A es una matriz de $n \times n$, entonces $A - A^T$ es antisimétrica.

43. **(a)** Demuestre que cualquier matriz cuadrada A puede escribirse como la suma de una matriz simétrica y una matriz antisimétrica. [*Sugerencia:* considere el Teorema 3.5 y el ejercicio 42.]
(b) Ejemplifique el inciso (a) con la matriz

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

La ***traza*** *de una matriz* $A = [a_{ij}]$ *de* $n \times n$ *es la suma de las entradas en su diagonal principal y se denota mediante* $\text{tr}(A)$. *Esto es,*

$$\text{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

44. Si A y B son matrices de $n \times n$, demuestre las siguientes propiedades de la traza:
(a) $\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B)$
(b) $\text{tr}(kA) = k\text{tr}(A)$, donde k es un escalar

45. Demuestre que si A y B son matrices de $n \times n$, entonces $\text{tr}(AB) = \text{tr}(BA)$.

46. Si A es cualquier matriz, ¿a qué es igual $\text{tr}(AA^T)$?

47. Demuestre que no hay matrices de A y B de 2×2 tales que $AB - BA = I_2$.

3.3 La inversa de una matriz

En esta sección se regresa a la descripción matricial $A\mathbf{x} = \mathbf{b}$ de un sistema de ecuaciones lineales y se buscan formas de usar el álgebra matricial para resolver el sistema. Por analogía, considere la ecuación $ax = b$, donde a, b y x representan números reales y se quiere resolver para x. Rápidamente puede suponer que se quiere $x = b/a$ como la solución, pero debe recordar que es cierto sólo si $a \neq 0$. Al proceder con más lentitud y suponer que $a \neq 0$, se llegará a la solución mediante la siguiente secuencia de pasos:

$$ax = b \Rightarrow \frac{1}{a}(ax) = \frac{1}{a}(b) \Rightarrow \left(\frac{1}{a}(a)\right)x = \frac{b}{a} \Rightarrow 1 \cdot x = \frac{b}{a} \Rightarrow x = \frac{b}{a}$$

(¡Este ejemplo muestra cuánto se hace en la cabeza y cuántas propiedades de aritmética y álgebra se dan por sentadas!)

Para imitar este procedimiento para la ecuación matricial $A\mathbf{x} = \mathbf{b}$, ¿qué se necesita? Se necesita encontrar una matriz A' (análoga a $1/a$) tal que $A'A = I$, una matriz identidad (análoga a 1). *Si* tal matriz existe (análoga al requisito de que $a \neq 0$), entonces puede hacer la siguiente secuencia de cálculos:

$$A\mathbf{x} = \mathbf{b} \Rightarrow A'(A\mathbf{x}) = A'\mathbf{b} \Rightarrow (A'A)\mathbf{x} = A'\mathbf{b} \Rightarrow I\mathbf{x} = A'\mathbf{b} \Rightarrow \mathbf{x} = A'\mathbf{b}$$



(¿Por qué se justificaría cada uno de estos pasos?)

La meta en esta sección es determinar precisamente cuándo es posible encontrar tal matriz A'. De hecho, se insistirá un poco más: se quiere no sólo $A'A = I$, sino también $AA' = I$. Este requisito fuerza a A y A' a ser matrices cuadradas. (¿Por qué?)

Definición Si A es una matriz de $n \times n$, una ***inversa*** de A es una matriz A' de $n \times n$ con la siguiente propiedad

$$AA' = I \quad \text{y} \quad A'A = I$$

donde $I = I_n$ es la matriz identidad. Si tal A' existe, entonces A es ***invertible***.

Ejemplo 3.22

Si $A = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$, entonces $A' = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$ es una inversa de A, pues

$$AA' = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad A'A = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Ejemplo 3.23

Demuestre que las siguientes matrices no son invertibles:

(a) $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ (b) $B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$

Solución

(a) Es fácil ver que la matriz O no tiene una inversa. Si lo tuviera, entonces sería una matriz O' tal que $OO' = I = O'O$. Pero el producto de la matriz cero con cualquiera otra matriz es la matriz cero y por tanto OO' nunca podría ser igual a la matriz identidad I.

(Note que esta prueba no hace alusión al tamaño de las matrices y por tanto es verdadera para matrices de $n \times n$ en general.)

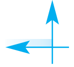
(b) Suponga que B no tiene inversa $B' = \begin{bmatrix} w & x \\ y & z \end{bmatrix}$. La ecuación $BB' = I$ produce

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

de donde se obtienen las ecuaciones

$$\begin{aligned} w \quad\quad + 2y \quad\quad &= 1 \\ x \quad\quad + 2z &= 0 \\ 2w \quad\quad + 4y \quad\quad &= 0 \\ 2x \quad\quad + 4z &= 1 \end{aligned}$$

Al restar dos veces la primera ecuación de la tercera produce $0 = -2$, que claramente es un absurdo. Por tanto, no hay solución. (La reducción por renglones produce el mismo resultado pero en realidad no es necesaria aquí.) Se deduce que no existe tal matriz B'; esto es: B no es invertible. (De hecho, incluso no tiene una inversa que funcione en un lado, ¡mucho menos en dos!)



Comentarios

- Aun cuando se vio que la multiplicación de matrices, en general, no es conmutativa, A' (si existe) debe satisfacer $A'A = AA'$.
- Los ejemplos anteriores plantean dos preguntas que se responderán en esta sección:

(1) ¿Cómo puede saberse cuándo una matriz tiene inversa?
(2) Si una matriz tiene inversa, ¿cómo puede encontrarla?

- No se reglamentó la posibilidad de que una matriz A pueda tener más de una inversa. El siguiente teorema garantiza que esto no puede suceder.

Teorema 3.6 Si A es una matriz invertible, entonces su inversa es única.

Demostración En matemáticas, una forma estándar de demostrar que sólo hay uno de algo es demostrar que no puede haber más de uno. De este modo, suponga que A tiene dos inversas, por ejemplo, A' y A''. Entonces

$$AA' = I = A'A \quad \text{y} \quad AA'' = I = A''A$$

Por tanto, $$A' = A'I = A'(AA'') = (A'A)A'' = IA'' = A''$$

En consecuencia, $A' = A''$, y la inversa es única.

Gracias a este teorema, ahora puede referirse a la inversa de una matriz invertible. A partir de ahora, cuando A sea invertible, su inversa (única) se denotará mediante A^{-1} (pronúnciese "A inversa").

Advertencia ¡No escriba $A^{-1} = \dfrac{1}{A}$! No existe una operación tal como la "división entre una matriz". Incluso si la hubiera, ¿cómo se podría dividir el *escalar* 1 entre la *matriz A*? Si alguna vez se siente tentado a "dividir" entre una matriz, lo que realmente quiere hacer es multiplicar por su inversa.

Ahora puede completar la analogía que se estableció al comienzo de esta sección.

Teorema 3.7 Si A es una matriz invertible de $n \times n$, entonces el sistema de ecuaciones lineales dado por $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$ para cualquier $\mathbf{b}$ en $\mathbb{R}^n$.

Demostración El Teorema 3.7 en esencia formaliza el comentario hecho al comienzo de esta sección. Se revisará nuevamente, ahora con un poco más de cuidado. Se pidió probar dos cosas: que $A\mathbf{x} = \mathbf{b}$ *tiene* una solución y que tiene *una sola* solución. (En matemáticas, tal demostración se conoce como demostración de "existencia y unicidad".)

Para demostrar que existe una solución, sólo necesita verificar que funciona $\mathbf{x} = A^{-1}\mathbf{b}$. Se comprueba que

$$A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$$

De modo que $A^{-1}\mathbf{b}$ satisface la ecuación $A\mathbf{x} = \mathbf{b}$ y por tanto existe *al menos* esta solución.

Para demostrar que esta solución es única, suponga que $\mathbf{y}$ es otra solución. Entonces $A\mathbf{y} = \mathbf{b}$, y al multiplicar ambos lados de la ecuación por A^{-1} a la izquierda, se obtiene la cadena de implicaciones

$$A^{-1}(A\mathbf{y}) = A^{-1}\mathbf{b} \Rightarrow (A^{-1}A)\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow I\mathbf{y} = A^{-1}\mathbf{b} \Rightarrow \mathbf{y} = A^{-1}\mathbf{b}$$

Por tanto, $\mathbf{y}$ es la misma solución que antes, y en consecuencia la solución es única.

De este modo, al regresar a las preguntas planteadas en los comentarios previos al Teorema 3.6, ¿cómo podría decir si una matriz es invertible y cómo puede encontrar su inversa cuando es invertible? Dentro de poco se dará un procedimiento general, pero la situación para las de matrices de 2×2 es suficientemente simple para garantizar que es singular.

Teorema 3.8 Si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, entonces A es invertible si $ad - bc \neq 0$, en cuyo caso

$$A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Si $ad - bc = 0$, entonces A no es invertible.

La expresión $ad - bc$ se llama ***determinante*** de A y se denota det A. Por tanto, la fórmula para la inversa de $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ (cuando existe) es $\dfrac{1}{\det A}$ multiplicado por la matriz obtenida al intercambiar las entradas en la diagonal principal y cambiar los signos en las otras dos entradas. Además de dar esta fórmula, el Teorema 3.8 dice que una matriz A de 2×2 es invertible si y sólo si det $A \neq 0$. En el capítulo 4 se verá que el determinante puede definirse para todas las matrices cuadradas y que este resultado sigue siendo cierto, aunque no hay una fórmula simple para la inversa de matrices cuadradas más grandes.

Demostración Suponga que $\det A = ad - bc \neq 0$. Entonces

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \begin{bmatrix} ad - bc & -ab + ba \\ cd - dc & -cb + da \end{bmatrix} = \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} = \det A\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

De igual modo,

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \det A\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Dado que $\det A \neq 0$, puede multiplicar ambos lados de cada ecuación por $1/\det A$ para obtener

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\left(\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

y

$$\left(\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\right)\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

[Note que se usó la propiedad (d) del Teorema 3.3.] Por tanto, la matriz

$$\frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

satisface la definición de una inversa, de modo que A es invertible. Puesto que la inversa de A es única, por el Teorema 3.6, debe tener

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Por el contrario, suponga que $ad - bc = 0$. Se considerarán por separado los casos donde $a \neq 0$ y donde $a = 0$. Si $a \neq 0$, entonces $d = bc/a$, de modo que la matriz puede escribirse como

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ ac/a & bc/a \end{bmatrix} = \begin{bmatrix} a & b \\ ka & kb \end{bmatrix}$$

donde $k = c/a$. En otras palabras, el segundo renglón de A es un múltiplo del primero. Al referirse al ejemplo 3.23(b), se ve que si A tiene una inversa $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$, entonces

$$\begin{bmatrix} a & b \\ ka & kb \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

y el sistema correspondiente de ecuaciones lineales

$$\begin{aligned} aw \qquad + by \qquad &= 1 \\ ax \qquad + bz &= 0 \\ kaw \qquad + kby \qquad &= 0 \\ kax \qquad + kbz &= 1 \end{aligned}$$



no tiene solución. (¿Por qué?)

Si $a = 0$, entonces $ad - bc = 0$ implica que $bc = 0$, y por tanto b o c es 0. En consecuencia, A es de la forma

$$\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}$$

En el primer caso, $\begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ * & * \end{bmatrix} \neq \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. De igual modo, $\begin{bmatrix} 0 & b \\ 0 & d \end{bmatrix}$ no



puede tener una inversa. (Verifique esto.)

En consecuencia, si $ad - bc = 0$, entonces A no es invertible.

Ejemplo 3.24

Encuentre las inversas de $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ y $B = \begin{bmatrix} 12 & -15 \\ 4 & -5 \end{bmatrix}$, si existen.

Solución Se tiene det $A = 1(4) - 2(3) = -2 \neq 0$, de modo que A es invertible, con

$$A^{-1} = \frac{1}{-2}\begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$$

(Compruebe esto.)

Por otra parte, $B = 12(-5) - (-15)(4) = 0$, de modo que B no es invertible.

Ejemplo 3.25

Use la inversa de la matriz de coeficientes para resolver el sistema lineal

$$\begin{aligned} x + 2y &= 3 \\ 3x + 4y &= -2 \end{aligned}$$

Solución La matriz de coeficientes es la matriz $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, cuya inversa se calculó en el ejemplo 3.24. Por el Teorema 3.7, $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$. Aquí se tiene $\mathbf{b} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$; por tanto, la solución al sistema dado es

$$\mathbf{x} = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -8 \\ \frac{11}{2} \end{bmatrix}$$

Comentario Resolver un sistema lineal $A\mathbf{x} = \mathbf{b}$ mediante $\mathbf{x} = A^{-1}\mathbf{b}$ parecería ser un buen método. Por desgracia, excepto para matrices de coeficientes de 2×2 y matrices con ciertas formas especiales, casi siempre es más rápido usar eliminación gaussiana o de Gauss-Jordan para encontrar directamente la solución. (Vea el ejercicio 13.) Más aún, la técnica del ejemplo 3.25 funciona sólo cuando la matriz de coeficientes es cuadrada e invertible, mientras que los métodos de eliminación siempre pueden aplicarse.

Propiedades de las matrices invertibles

El siguiente teorema registra algunas de las propiedades más importantes de las matrices invertibles.

Teorema 3.9

a. Si A es una matriz invertible, entonces A^{-1} es invertible y

$$(A^{-1})^{-1} = A$$

b. Si A es una matriz invertible y c es un escalar distinto de cero, entonces cA es una matriz invertible y

$$(cA)^{-1} = \frac{1}{c}A^{-1}$$

c. Si A y B son matrices invertibles del mismo tamaño, entonces AB es invertible y

$$(AB)^{-1} = B^{-1}A^{-1}$$

d. Si A es una matriz invertible, entonces A^T es invertible y

$$(A^T)^{-1} = (A^{-1})^T$$

e. Si A es una matriz invertible, entonces A^n es invertible para todo entero n no negativo y

$$(A^n)^{-1} = (A^{-1})^n$$

Demostración Se demostrarán las propiedades (a), (c) y (e), y las propiedades (b) y (d) se dejarán para su demostración en los ejercicios 14 y 15.

(a) Para demostrar que A^{-1} es invertible, debe argumentar que existe una matriz X tal que

$$A^{-1}X = I = XA^{-1}$$

Pero ciertamente A satisface estas ecuaciones en lugar de X, de modo que A^{-1} es invertible y A es *una* inversa de A^{-1}. Dado que las inversas son únicas, esto significa que $(A^{-1})^{-1} = A$.

(c) Aquí debe demostrar que existe una matriz X tal que

$$(AB)X = I = X(AB)$$

La afirmación es que la sustitución de $B^{-1}A^{-1}$ por X funciona. Se comprueba que

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

donde se usó asociatividad para recorrer los paréntesis. De igual modo, $(B^{-1}A^{-1})(AB) = I$ (¡compruébelo!), de modo que AB es invertible y su inversa es $B^{-1}A^{-1}$.

(e) La idea básica aquí es suficientemente sencilla. Por ejemplo, cuando $n = 2$, se tiene

$$A^2(A^{-1})^2 = AAA^{-1}A^{-1} = AIA^{-1} = AA^{-1} = I$$

De igual modo, $(A^{-1})^2A^2 = I$. Por ende, $(A^{-1})^2$ es la inversa de A^2. No es difícil ver que un argumento similar funciona para cualquier valor entero superior de n. Sin embargo, la inducción matemática es la forma de realizar la demostración.

El paso básico es cuando $n = 0$, en cuyo caso se pedirá probar que A^0 es invertible y que

$$(A^0)^{-1} = (A^{-1})^0$$

Esto es lo mismo que demostrar que I es invertible y que $I^{-1} = I$, lo que claramente es verdadero. (¿Por qué? Vea el ejercicio 16.)

Ahora suponga que el resultado es verdadero cuando $n = k$, donde k es un entero no negativo específico. Esto es: la hipótesis de inducción es suponer que A^k es invertible y que

$$(A^k)^{-1} = (A^{-1})^k$$

El paso de inducción requiere demostrar que A^{k+1} es invertible y que $(A^{k+1})^{-1} = (A^{-1})^{k+1}$. Ahora se sabe de (c) que $A^{k+1} = A^kA$ es invertible, pues A y (por hipótesis) A^k son ambos invertibles. Más aún,

$$\begin{aligned}
(A^{-1})^{k+1} &= (A^{-1})^kA^{-1} \\
&= (A^k)^{-1}A^{-1} && \text{por la hipótesis de inducción} \\
&= (AA^k)^{-1} && \text{por la propiedad (c)} \\
&= (A^{k+1})^{-1}
\end{aligned}$$

Por tanto, A^n es invertible para todo entero no negativo n y $(A^n)^{-1} = (A^{-1})^n$ por el principio de inducción matemática.

Comentarios

- Aunque todas las propiedades del Teorema 3.9 son útiles, (c) es la que se debe destacar. Acaso sea la más importante propiedad algebraica de las inversas de matrices. También es la que es más fácil de que salga mal. En el ejercicio 17 se le pedirá proporcionar un contraejemplo para demostrar que, contrario a lo que se quisiera, $(AB)^{-1} \neq A^{-1}B^{-1}$ en general. La propiedad correcta, $(AB)^{-1} = B^{-1}A^{-1}$, en ocasiones se conoce como la regla calcetas-zapatos, porque, aunque uno se pone las calcetas antes que los zapatos, se los quita en el orden inverso.
- La propiedad (c) generaliza los productos de un finito de matrices invertibles: si $A_1, A_2, \ldots, A_n$ son matrices invertibles del mismo tamaño, entonces $A_1A_2 \cdots A_n$ es invertible y

$$(A_1A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$$

(Vea el ejercicio 18.) Por tanto, puede afirmarse:

> El inverso de un producto de matrices invertibles es el producto de sus inversas en el orden inverso.

- Dado que, para números reales, $\dfrac{1}{a+b} \neq \dfrac{1}{a} + \dfrac{1}{b}$, no debe esperar que para matrices cuadradas, $(A + B)^{-1} = A^{-1} + B^{-1}$ (y, de hecho, esto no es cierto en general; vea el ejercicio 19). De hecho, excepto para matrices especiales, no hay fórmula para $(A + B)^{-1}$.
- La propiedad (e) permite definir potencias enteras negativas de una matriz invertible:

> Si A es una matriz invertible y n es un entero positivo, entonces A^{-n} se define por
>
> $$A^{-n} = (A^{-1})^n = (A^n)^{-1}$$

Con esta definición, puede demostrarse que las reglas para exponenciación, $A^rA^s = A^{r+s}$ y $(A^r)^s = A^{rs}$, se sostienen para todos los enteros r y s, siempre que A sea invertible.

Un uso de las propiedades algebraicas de las matrices es ayudar a resolver ecuaciones que involucran matrices. El siguiente ejemplo ilustra el proceso. Note que debe poner particular atención al orden de las matrices en el producto.

Ejemplo 3.26

Resuelva la siguiente ecuación matricial para encontrar X (suponga que las matrices involucradas son tales que todas las operaciones indicadas están definidas):

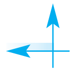
$$A^{-1}(BX)^{-1} = (A^{-1}B^3)^2$$

Solución Existen muchas formas de proceder aquí. Una solución es

$$\begin{aligned}
A^{-1}(BX)^{-1} = (A^{-1}B^3)^2 &\Rightarrow ((BX)A)^{-1} = (A^{-1}B^3)^2 \\
&\Rightarrow [((BX)A)^{-1}]^{-1} = [(A^{-1}B^3)^2]^{-1} \\
&\Rightarrow (BX)A = [(A^{-1}B^3)(A^{-1}B^3)]^{-1} \\
&\Rightarrow (BX)A = B^{-3}(A^{-1})^{-1}B^{-3}(A^{-1})^{-1} \\
&\Rightarrow BXA = B^{-3}AB^{-3}A \\
&\Rightarrow B^{-1}BXAA^{-1} = B^{-1}B^{-3}AB^{-3}AA^{-1} \\
&\Rightarrow IXI = B^{-4}AB^{-3}I \\
&\Rightarrow X = B^{-4}AB^{-3}
\end{aligned}$$



(¿Puede justificar cada paso?) Note el uso cuidadoso del Teorema 3.9(c) y el desarrollo de $(A^{-1}B^3)^2$. También se usó libremente la asociatividad de la multiplicación de matrices para simplificar la colocación (o eliminación) de los paréntesis.



Matrices elementales

La multiplicación de matrices se usará para dar una perspectiva diferente a la reducción de matrices por renglones. En el proceso, descubrirá muchos nuevos e importantes conocimientos de la naturaleza de las matrices invertibles.

Si

$$E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 5 & 7 \\ -1 & 0 \\ 8 & 3 \end{bmatrix}$$

se encuentra que

$$EA = \begin{bmatrix} 5 & 7 \\ 8 & 3 \\ -1 & 0 \end{bmatrix}$$

En otras palabras, multiplicar A por E (a la izquierda) tiene el mismo efecto que intercambiar los renglones 2 y 3 de A. ¿Qué es lo significativo de E? Simplemente es la matriz que se obtiene al aplicar la misma operación elemental con renglones, $R_2 \leftrightarrow R_3$, a la matriz identidad I_3. Es evidente que esto siempre funciona.

Definición Una ***matriz elemental*** es aquella matriz que puede obtenerse al realizar una operación elemental con renglones sobre una matriz identidad.

Dado que existen tres tipos de operaciones elementales con renglones, existen tres tipos correspondientes de matrices elementales. He aquí algunas de las matrices más elementales.

Ejemplo 3.27 Sean

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -2 & 0 & 1 \end{bmatrix}$$

Cada una de esas matrices se obtuvo a partir de la matriz identidad I_4 al aplicar una sola operación elemental con renglones. La matriz E_1 corresponde a $3R_2$, E_2 a $R_1 \leftrightarrow R_3$ y E_3 a $R_4 - 2R_2$. Observe que cuando se multiplica por la izquierda una matriz de $4 \times n$ por una de tales matrices elementales, la correspondiente operación elemental con renglones se realiza sobre la matriz. Por ejemplo, si

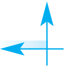
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}$$

entonces

$$E_1A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 3a_{21} & 3a_{22} & 3a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix}, \quad E_2A = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \\ a_{41} & a_{42} & a_{43} \end{bmatrix},$$

y

$$E_3A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} - 2a_{21} & a_{42} - 2a_{22} & a_{43} - 2a_{23} \end{bmatrix}$$



El ejemplo 3.27 y los ejercicios 24-30 deben convencerlo de que *cualquier* operación elemental con renglones sobre *cualquier* matriz puede lograrse al multiplicar por la izquierda por una matriz elemental adecuada. Este hecho se registra como teorema, del que se omite la demostración.

Teorema 3.10 Sea E la matriz elemental que se obtiene al realizar una operación elemental con renglones sobre I_n. Si la misma operación elemental con renglones se realiza sobre una matriz A de $n \times n$, el resultado es el mismo que la matriz EA.

Comentario Desde un punto de vista computacional, no es buena idea usar matrices elementales para realizar operaciones elementales con renglones, sólo hágalas directamente. Sin embargo, las matrices elementales pueden proporcionar algunas intuiciones valiosas acerca de las matrices invertibles y la solución de sistemas de ecuaciones lineales.

Ya se observó que toda operación elemental con renglones puede "deshacerse" o "invertirse". Esta misma observación aplicada a matrices elementales muestra que son invertibles.

Ejemplo 3.28 Sean

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$



Entonces E_1 corresponde a $R_2 \leftrightarrow R_3$, que se deshace al efectuar $R_2 \leftrightarrow R_3$ nuevamente. Por tanto, $E_1^{-1} = E_1$. (Compruebe al demostrar que $E_1^2 = E_1E_1 = I$.) La matriz E_2 proviene de $4R_2$, que se deshace al aplicar $\frac{1}{4}R_2$. Por tanto,

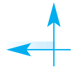
$$E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

que puede comprobarse con facilidad. Finalmente, E_3 corresponde a la operación elemental con renglones $R_3 - 2R_1$, que puede deshacerse mediante la operación elemental con renglones $R_3 + 2R_1$. De modo que, en este caso,

$$E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$$



(Nuevamente, es fácil comprobar esto al confirmar que el producto de esta matriz y E_3, en ambos órdenes, es I.)



Note que no sólo cada matriz elemental es invertible, sino que su inversa es otra matriz elemental del mismo tipo. Este hallazgo se registra como el siguiente teorema.

Teorema 3.11 Cada matriz elemental es invertible y su inversa es una matriz elemental del mismo tipo.

El teorema fundamental de las matrices invertibles

Ahora está en posición de probar uno de los principales resultados de este libro: un conjunto de caracterizaciones equivalentes de lo que significa para una matriz ser invertible. En un sentido, gran parte del álgebra lineal está conectada a este teorema, ya sea en el desarrollo de dichas caracterizaciones o en su aplicación. Como puede esperar, dada esta introducción, este teorema se usará bastante. ¡Hágalo su amigo!

Al Teorema 3.12 se le conocerá como la primera versión del teorema fundamental, pues se le harán agregados en capítulos posteriores. Se le recuerda que, cuando se dice que un conjunto de enunciados acerca de una matriz A son equivalentes, se entiende que, para una A dada, los enunciados son todos verdaderos o todos falsos.

Teorema 3.12 **El teorema fundamental de las matrices invertibles: versión 1**

Sea A una matriz de $n \times n$. Los siguientes enunciados son equivalentes:

a. A es invertible.
b. $A\mathbf{x} = \mathbf{b}$ tiene una solución única para todo $\mathbf{b}$ en $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial.
d. La forma escalonada reducida por renglones de A es I_n.
e. A es un producto de matrices elementales.

Demostración El teorema se establecerá al demostrar la cadena circular de implicaciones

$$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (a)$$

(a) $\Rightarrow$ (b) Ya se demostró que si A es invertible, entonces $A\mathbf{x} = \mathbf{b}$ tiene la solución única $\mathbf{x} = A^{-1}\mathbf{b}$ para cualquier $\mathbf{b}$ en $\mathbb{R}^n$ (Teorema 3.7).

(b) $\Rightarrow$ (c) Suponga que $A\mathbf{x} = \mathbf{b}$ tiene una solución única para cualquier $\mathbf{b}$ en $\mathbb{R}^n$. Esto implica, en particular, que $A\mathbf{x} = \mathbf{0}$ tiene una solución única. Pero un sistema homogéneo $A\mathbf{x} = \mathbf{0}$ siempre tiene $\mathbf{x} = \mathbf{0}$ como *una* solución. Así que, en este caso, $\mathbf{x} = \mathbf{0}$ debe ser *la* solución.

(c) $\Rightarrow$ (d) Suponga que $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial. El correspondiente sistema de ecuaciones es

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0 \end{aligned}$$

y se supone que esta solución es

$$\begin{aligned} x_1 \qquad\qquad\qquad &= 0 \\ x_2 \qquad\quad &= 0 \\ \ddots \qquad &\\ x_n &= 0 \end{aligned}$$

En otras palabras, la eliminación de Gauss-Jordan aplicada a la matriz aumentada del sistema produce

$$[A|\mathbf{0}] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{array}\right] = [I_n|\mathbf{0}]$$

Por tanto, la forma escalonada reducida por renglones de A es I_n.

(d) $\Rightarrow$ (e) Si supone que la forma escalonada reducida por renglones de A es I_n, entonces A puede reducirse a I_n usando una secuencia finita de operaciones elementales con renglones. Por el Teorema 3.10, cada una de dichas operaciones elementales con renglones puede lograrse al multiplicar la izquierda por una matriz elemental adecuada. Si la secuencia adecuada de matrices elementales es $E_1, E_2, \ldots, E_k$ (en ese orden), entonces se tiene

$$E_k \cdots E_2E_1A = I_n$$

De acuerdo con el Teorema 3.11, dichas matrices elementales son todas invertibles. Por tanto, también lo es su producto, y se tiene

$$A = (E_k \cdots E_2E_1)^{-1}I_n = (E_k \cdots E_2E_1)^{-1} = E_1^{-1}E_2^{-1} \cdots E_k^{-1}$$

De nuevo, cada E_i^{-1} es otra matriz elemental, por el Teorema 3.11, así que A se escribió como un producto de matrices elementales, como se requería.

(e) $\Rightarrow$ (a) si A es un producto de matrices elementales, entonces A es invertible, pues las matrices elementales son invertibles y los productos de las matrices invertibles son invertibles.

Ejemplo 3.29 Si es posible, exprese $A = \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix}$ como un producto de matrices elementales.

Solución Se reduce A por renglones del modo siguiente:

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 3 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 3 \\ 2 & 3 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 3 \\ 0 & -3 \end{bmatrix}$$
$$\xrightarrow{R_1 + R_2} \begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix} \xrightarrow{-\frac{1}{3}R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2$$

Por tanto, la forma escalonada reducida por renglones de A es la matriz identidad, así que el teorema fundamental asegura que A es invertible y puede escribirse como un producto de matrices elementales. Se tiene $E_4E_3E_2E_1A = I$, donde

$$E_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad E_4 = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{3} \end{bmatrix}$$

son las matrices elementales correspondientes a las cuatro operaciones elementales con renglones utilizadas para reducir A a I. Como en la prueba del teorema, se tiene

$$A = (E_4E_3E_2E_1)^{-1} = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -3 \end{bmatrix}$$

como se requiere.

Comentario Puesto que la secuencia de operaciones elementales con renglones que transforma A en I no es única, tampoco lo es la representación de A como producto de matrices elementales. (Encuentre una forma diferente de expresar A como un producto de matrices elementales).

"Nunca lleve un cañón al escenario en el primer acto, a menos que pretenda dispararlo en el último acto." –Anton Chejov

El teorema fundamental es sorprendentemente poderoso. Para ilustrar su poder, considere dos de sus consecuencias. La primera es que, aunque la definición de una matriz invertible afirma que una matriz A es invertible si existe una matriz B tal que *tanto* $AB = I$ *como* $BA = I$ se satisfacen, sólo es necesario comprobar *una* de dichas ecuaciones. Por tanto, ¡puede reducirse el trabajo a la mitad!

Teorema 3.13 Sea A una matriz cuadrada. Si B es una matriz cuadrada tal que $AB = I$ o $BA = I$, entonces A es invertible y $B = A^{-1}$.

Demostración Suponga $BA = I$. Considere la ecuación $A\mathbf{x} = \mathbf{0}$. Al multiplicar por la izquierda por B, se tiene $BA\mathbf{x} = B\mathbf{0}$. Esto implica que $\mathbf{x} = I\mathbf{x} = \mathbf{0}$. Por tanto, el sistema representado por $A\mathbf{x} = \mathbf{0}$ tiene la solución única $\mathbf{x} = \mathbf{0}$. A partir de la equivalencia de (c) y (a) en el teorema fundamental, se sabe que A es invertible. (Esto es: existe A^{-1} y satisface $AA^{-1} = I = A^{-1}A$.)

Si ahora multiplica por la derecha ambos lados de $BA = I$ por A^{-1}, obtiene

$$BAA^{-1} = IA^{-1} \Rightarrow BI = A^{-1} \Rightarrow B = A^{-1}$$

(La demostración en el caso de $AB = I$ se deja como ejercicio 41.)

La siguiente consecuencia del teorema fundamental es la base de un método eficiente para calcular la inversa de una matriz.

Teorema 3.14 Sea A una matriz cuadrada. Si una secuencia de operaciones elementales con renglones reduce A a I, entonces la misma secuencia de operaciones elementales transforma I en A^{-1}.

Demostración Si A es equivalente por renglones a I, entonces puede lograrse la reducción mediante multiplicación por la izquierda por una secuencia $E_1, E_2, \ldots, E_k$ de matrices elementales. Por tanto, se tiene $E_k \cdots E_2E_1A = I$. Al hacer $B = E_k \cdots E_2E_1$ se obtiene $BA = I$. Por el Teorema 3.13, A es invertible y $A^{-1} = B$. Ahora, aplicar la misma secuencia de operaciones elementales con renglones a I equivale a multiplicar por la izquierda I por $E_k \cdots E_2E_1 = B$. El resultado es

$$E_k \cdots E_2E_1I = BI = B = A^{-1}$$

Por tanto, I se transforma en A^{-1} mediante la misma secuencia de operaciones elementales con renglones.

El método de Gauss-Jordan para calcular la inversa

Es posible realizar operaciones con renglones sobre A e I simultáneamente al construir una "matriz superaumentada" $[A \mid I]$. El Teorema 3.14 muestra que si A es equivalente por renglones a I [lo cual, por el teorema fundamental (d) $\Leftrightarrow$ (a), significa que A es invertible], entonces operaciones elementales con renglones producirán

$$[A \mid I] \longrightarrow [I \mid A^{-1}]$$

Si A no puede reducirse a I, entonces el teorema fundamental garantiza que A no es invertible.

El procedimiento recién descrito es simplemente la eliminación de Gauss-Jordan efectuada sobre una matriz aumentada de $n \times 2n$, en lugar de sobre una de $n \times (n + 1)$. Otra forma de ver este procedimiento es observar el problema de encontrar A^{-1} al resolver la ecuación matricial $AX = I_n$ para una matriz X de $n \times n$. (Esto es suficiente, por el teorema fundamental, pues una inversa correcta de A debe ser una inversa de dos lados.) Si las columnas de X se denotan $\mathbf{x}_1, \ldots, \mathbf{x}_n$, entonces esta ecuación matricial es equivalente a resolver para las columnas de X, una a la vez. Dado que las columnas de I_n son los vectores unitarios estándar $\mathbf{e}_1, \ldots, \mathbf{e}_n$, entonces se tienen n sistemas de ecuaciones lineales, todas con matriz de coeficientes A:

$$A\mathbf{x}_1 = \mathbf{e}_1, \ldots, A\mathbf{x}_n = \mathbf{e}_n$$

Dado que se necesita la misma secuencia de operaciones con renglones para llevar A a la forma escalonada reducida por renglones en cada caso, las matrices aumentadas para estos sistemas, $[A \mid \mathbf{e}_1], \ldots, [A \mid \mathbf{e}_n]$, pueden combinarse como

$$[A \mid \mathbf{e}_1\, \mathbf{e}_2 \cdots \mathbf{e}_n] = [A \mid I_n]$$

Ahora se aplican operaciones con renglones para tratar de reducir A a I_n, lo cual, si tiene éxito, resolverá simultáneamente para las columnas de A^{-1}, lo que transformará I_n en A^{-1}.

Este uso de la eliminación de Gauss-Jordan se ilustra con tres ejemplos.

Ejemplo 3.30 Encuentre la inversa de

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix}$$

si existe.

Solución La eliminación de Gauss-Jordan produce

$$[A\,|\,I] = \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 2 & 2 & 4 & 0 & 1 & 0 \\ 1 & 3 & -3 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3-R_1]{R_2-2R_1} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & -2 & 6 & -2 & 1 & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{(-\frac{1}{2})R_2} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -2 & -1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3-R_2} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

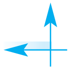
$$\xrightarrow[R_2+3R_3]{R_1+R_3} \left[\begin{array}{rrr|rrr} 1 & 2 & 0 & -1 & \frac{1}{2} & 1 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

$$\xrightarrow{R_1-2R_2} \left[\begin{array}{rrr|rrr} 1 & 0 & 0 & 9 & -\frac{3}{2} & -5 \\ 0 & 1 & 0 & -5 & 1 & 3 \\ 0 & 0 & 1 & -2 & \frac{1}{2} & 1 \end{array}\right]$$

Por tanto,

$$A^{-1} = \begin{bmatrix} 9 & -\frac{3}{2} & -5 \\ -5 & 1 & 3 \\ -2 & \frac{1}{2} & 1 \end{bmatrix}$$

(Siempre debe comprobar que $AA^{-1} = I$ mediante multiplicación directa. Por el Teorema 3.13, no necesita comprobar que también $A^{-1}A = I$.)



Comentario Note que se usó la variante de eliminación de Gauss-Jordan que primero introduce todos los ceros *abajo* de los 1 pivote, de izquierda a derecha y de arriba abajo, y luego crea ceros *arriba* de los 1 pivote, de derecha a izquierda y de abajo arriba. Este planteamiento ahorra cálculos, como se anotó en el capítulo 2, pero puede encontrarlo más sencillo, cuando trabaja a mano, para crear *todos* los ceros en cada columna conforme avance. Desde luego, la respuesta será la misma.

Ejemplo 3.31 Encuentre la inversa de

$$A = \begin{bmatrix} 2 & 1 & -4 \\ -4 & -1 & 6 \\ -2 & 2 & -2 \end{bmatrix}$$

si existe.

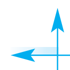
Solución Proceda como en el ejemplo 3.30, adjunte la matriz identidad a A y luego trate de convertir $[A \mid I]$ en $[I \mid A^{-1}]$.

$$[A \mid I] = \left[\begin{array}{rrr|rrr} 2 & 1 & -4 & 1 & 0 & 0 \\ -4 & -1 & 6 & 0 & 1 & 0 \\ -2 & 2 & -2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow[R_3+R_1]{R_2+2R_1} \left[\begin{array}{rrr|rrr} 2 & 1 & -4 & 1 & 0 & 0 \\ 0 & 1 & -2 & 2 & 1 & 0 \\ 0 & 3 & -6 & 1 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_3-3R_2} \left[\begin{array}{rrr|rrr} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & -3 & 2 & 1 & 0 \\ 0 & 0 & 0 & -5 & -3 & 1 \end{array}\right]$$

En este punto se ve que no es posible reducir A a I, pues existe un renglón de ceros en el lado izquierdo de la matriz aumentada. En consecuencia, A no es invertible.



Como ilustra el ejemplo siguiente, todo funciona de la misma forma sobre $\mathbb{Z}_p$, donde p es primo.

Ejemplo 3.32 Encuentre la inversa de

$$A = \begin{bmatrix} 2 & 2 \\ 2 & 0 \end{bmatrix}$$

si existe, sobre $\mathbb{Z}_3$.

Solución 1 Use el método de Gauss-Jordan y recuerde que todos los cálculos son en $\mathbb{Z}_3$.

$$[A \mid I] = \left[\begin{array}{rr|rr} 2 & 2 & 1 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{2R_1} \left[\begin{array}{rr|rr} 1 & 1 & 2 & 0 \\ 2 & 0 & 0 & 1 \end{array}\right]$$

$$\xrightarrow{R_2+R_1} \left[\begin{array}{rr|rr} 1 & 1 & 2 & 0 \\ 0 & 1 & 2 & 1 \end{array}\right]$$

$$\xrightarrow{R_1+2R_2} \left[\begin{array}{rr|rr} 1 & 0 & 0 & 2 \\ 0 & 1 & 2 & 1 \end{array}\right]$$

Por tanto, $A^{-1} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}$ y es fácil comprobar que, sobre $\mathbb{Z}_3$, $AA^{-1} = I$.

Solución 2 Dado que A es una matriz de 2×2, también puede calcular A^{-1} usando la fórmula dada en el Teorema 3.8. El determinante de A es

$$\det A = 2(0) - 2(2) = -1 = 2$$

en $\mathbb{Z}_3$ (ya que $2 + 1 = 0$). Por tanto, A^{-1} existe y está dada por la fórmula del Teorema 3.8. Sin embargo, aquí debe tener cuidado, porque la fórmula introduce la "fracción" $1/\det A$ y en $\mathbb{Z}_3$ no existen fracciones. Debe usar inversos multiplicativos en lugar de división.

En lugar de $1/\det A = 1/2$, use 2^{-1}; esto es, encuentre el número x que satisfaga la ecuación $2x = 1$ en $\mathbb{Z}_3$. Es fácil ver que $x = 2$ es la solución que se quiere: en $\mathbb{Z}_3$, $2^{-1} = 2$, pues $2(2) = 1$. La fórmula para A^{-1} ahora se convierte en

$$A^{-1} = 2^{-1}\begin{bmatrix} 0 & -2 \\ -2 & 2 \end{bmatrix} = 2\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \end{bmatrix}$$

que concuerda con la solución anterior.

Ejercicios 3.3

En los ejercicios 1-10, encuentre la inversa de la matriz dada (si existe) usando el Teorema 3.8.

1. $\begin{bmatrix} 4 & 7 \\ 1 & 2 \end{bmatrix}$ **2.** $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

3. $\begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$ **4.** $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

5. $\begin{bmatrix} \frac{3}{4} & \frac{3}{5} \\ \frac{5}{6} & \frac{2}{3} \end{bmatrix}$ **6.** $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1\sqrt{2} \end{bmatrix}$

7. $\begin{bmatrix} -1.5 & -4.2 \\ 0.5 & 2.4 \end{bmatrix}$ **8.** $\begin{bmatrix} 3.55 & 0.25 \\ 8.52 & 0.60 \end{bmatrix}$

9. $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$

10. $\begin{bmatrix} 1/a & 1/b \\ 1/c & 1/d \end{bmatrix}$, donde ni a, b, c ni d son 0.

En los ejercicios 11 y 12, resuelva el sistema dado usando el método del ejemplo 3.25.

11. $2x + y = -1$
$5x + 3y = 2$

12. $x_1 - x_2 = 2$
$x_1 + 2x_2 = 5$

13. Sea $A = \begin{bmatrix} 1 & 2 \\ 2 & 6 \end{bmatrix}$, $\mathbf{b}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$, $\mathbf{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$, y $\mathbf{b}_3 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

(a) Encuentre A^{-1} y úselo para resolver los tres sistemas $A\mathbf{x} = \mathbf{b}_1$, $A\mathbf{x} = \mathbf{b}_2$ y $A\mathbf{x} = \mathbf{b}_3$.

(b) Resuelva los tres sistemas al mismo tiempo mediante la reducción por renglones de la matriz aumentada $[A \mid \mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3]$, usando la eliminación de Gauss-Jordan.

(c) Cuente cuidadosamente el número total de multiplicaciones individuales que realizó en (a) y en (b). Debe descubrir que, incluso para este ejemplo de 2×2, un método usa menos operaciones. Para sistemas más grandes, la diferencia es incluso más pronunciada, y esto explica por qué los sistemas de cómputo no usan uno de dichos métodos para resolver sistemas lineales.

14. Demuestre el Teorema 3.9(b).

15. Demuestre el Teorema 3.9(d).

16. Demuestre que la matriz identidad I_n de $n \times n$ es invertible y que $I_n^{-1} = I_n$.

17. (a) Proporcione un contraejemplo para demostrar que $(AB)^{-1} \neq A^{-1}B^{-1}$ en general.

(b) ¿En qué condiciones de A y B, $(AB)^{-1} = A^{-1}B^{-1}$? Pruebe su afirmación.

18. Mediante inducción, demuestre que si $A_1, A_2, \ldots, A_n$ son matrices invertibles del mismo tamaño, entonces el producto $A_1A_2 \cdots A_n$ es invertible y $(A_1A_2 \cdots A_n)^{-1} = A_n^{-1} \cdots A_2^{-1}A_1^{-1}$.

19. Proporcione un contraejemplo para demostrar que en general $(A + B)^{-1} \neq A^{-1} + B^{-1}$.

En los ejercicios 20-23, resuelva la ecuación matricial dada para X. Simplifique sus respuestas tanto como sea posible. (En palabras de Albert Einstein, "todo debe hacerse tan simple como sea posible, pero no más sencillo".) Suponga que todas las matrices son invertibles.

20. $XA^{-1} = A^3$ **21.** $AXB = (BA)^2$

22. $(A^{-1}X)^{-1} = (AB^{-1})^{-1}(AB^2)$ **23.** $ABXA^{-1}B^{-1} = I + A$

En los ejercicios 24-30, sean

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & -1 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 & -1 \\ 1 & 1 & 1 \\ 2 & 1 & -1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 2 & -1 \\ -3 & -1 & 3 \\ 2 & 1 & -1 \end{bmatrix}$$

En cada caso, encuentre una matriz elemental E que satisfaga la ecuación dada.

24. $EA = B$ **25.** $EB = A$ **26.** $EA = C$

27. $EC = A$ **28.** $EC = D$ **29.** $ED = C$

30. ¿Existe una matriz elemental E tal que $EA = D$? ¿Por qué sí o por qué no?

En los ejercicios 31-38, encuentre la inversa de la matriz elemental dada.

31. $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$ **32.** $\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

33. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **34.** $\begin{bmatrix} 1 & 0 \\ -\frac{1}{2} & 1 \end{bmatrix}$

35. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$ **36.** $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

37. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix}, c \neq 0$ **38.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix}, c \neq 0$

En los ejercicios 39 y 40, encuentre una secuencia de matrices elementales $E_1, E_2, \ldots, E_k$ tal que $E_k \cdots E_2 E_1 A = I$. Use esta secuencia para escribir A y A^{-1} como productos de matrices elementales.

39. $A = \begin{bmatrix} 1 & 0 \\ -1 & -2 \end{bmatrix}$ **40.** $A = \begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}$

41. Demuestre el Teorema 3.13 para el caso de $AB = I$.

42. **(a)** Demuestre que si A es invertible y $AB = O$, entonces $B = O$.

(b) Ofrezca un contraejemplo para demostrar que el resultado en el inciso (a) puede fallar si A no es invertible.

43. **(a)** Demuestre que si A es invertible y $BA = CA$, entonces $B = C$.

(b) Ofrezca un contraejemplo para demostrar que el resultado en el inciso (a) puede fallar si A no es invertible.

44. Una matriz cuadrada A se llama ***idempotente*** si $A^2 = A$. (La palabra *idempotente* viene de la palabra latina *idem*, que significa "igual", y *potere*, que significa "tener poder". Por tanto, algo que es idempotente tiene "la misma potencia" cuando se eleva al cuadrado.)

(a) Encuentre tres matrices de 2×2 idempotentes.

(b) Demuestre que la única matriz de $n \times n$ idempotente invertible es la matriz identidad.

45. Demuestre que si A es una matriz cuadrada que satisface la ecuación $A^2 - 2A + I = O$, entonces $A^{-1} = 2I - A$.

46. Demuestre que si una matriz simétrica es invertible, entonces su inversa también es simétrica.

47. Demuestre que si A y B son matrices cuadradas, y AB es invertible, entonces A y B son invertibles.

En los ejercicios 48-63, use el método de Gauss-Jordan para encontrar la inversa de la matriz dada (si existe).

48. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ **49.** $\begin{bmatrix} -2 & 4 \\ 3 & -1 \end{bmatrix}$

50. $\begin{bmatrix} 3 & -4 \\ -6 & 8 \end{bmatrix}$ **51.** $\begin{bmatrix} 1 & a \\ -a & 1 \end{bmatrix}$

52. $\begin{bmatrix} 2 & 0 & -1 \\ 1 & 5 & 1 \\ 2 & 3 & 0 \end{bmatrix}$ **53.** $\begin{bmatrix} 1 & -1 & 2 \\ 3 & 1 & 2 \\ 2 & 3 & -1 \end{bmatrix}$

54. $\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ **55.** $\begin{bmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{bmatrix}$

56. $\begin{bmatrix} 0 & a & 0 \\ b & 0 & c \\ 0 & d & 0 \end{bmatrix}$ **57.** $\begin{bmatrix} 0 & -1 & 1 & 0 \\ 2 & 1 & 0 & 2 \\ 1 & -1 & 3 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}$

58. $\begin{bmatrix} \sqrt{2} & 0 & 2\sqrt{2} & 0 \\ -4\sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 \end{bmatrix}$

59. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a & b & c & d \end{bmatrix}$ **60.** $\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_2$

61. $\begin{bmatrix} 4 & 2 \\ 3 & 4 \end{bmatrix}$ sobre $\mathbb{Z}_5$ **62.** $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 1 & 2 \\ 0 & 2 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_3$

63. $\begin{bmatrix} 1 & 5 & 0 \\ 1 & 2 & 4 \\ 3 & 6 & 1 \end{bmatrix}$ sobre $\mathbb{Z}_7$

Particionar matrices cuadradas grandes en ocasiones puede hacer que sus inversas sean más fáciles de calcular, en particular si los bloques tienen una forma manejable. En los ejercicios 64-68, verifique mediante multiplicación por bloque que la inversa de una matriz, si se particiona como se muestra, es como se afirma. (Suponga que todas las inversas existen como se necesita.)

64. $\begin{bmatrix} A & B \\ O & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & -A^{-1}BD^{-1} \\ O & D^{-1} \end{bmatrix}$

65. $\begin{bmatrix} O & B \\ C & I \end{bmatrix}^{-1} = \begin{bmatrix} -(BC)^{-1} & (BC)^{-1}B \\ C(BC)^{-1} & I - C(BC)^{-1}B \end{bmatrix}$

66. $\begin{bmatrix} I & B \\ C & I \end{bmatrix}^{-1} = \begin{bmatrix} (I - BC)^{-1} & -(I - BC)^{-1}B \\ -C(I - BC)^{-1} & I + C(I - BC)^{-1}B \end{bmatrix}$

67. $\begin{bmatrix} O & B \\ C & D \end{bmatrix}^{-1}$

$= \begin{bmatrix} -(BD^{-1}C)^{-1} & (BD^{-1}C)^{-1}BD^{-1} \\ D^{-1}C(BD^{-1}C)^{-1} & D^{-1} - D^{-1}C(BD^{-1}C)^{-1}BD^{-1} \end{bmatrix}$

68. $\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$, donde $P = (A - BD^{-1}C)^{-1}$, $Q = -PBD^{-1}$, $R = -D^{-1}CP$, y $S = D^{-1} + D^{-1}CPBD^{-1}$

En los ejercicios 69-72, particione la matriz dada de modo que pueda aplicar una de las fórmulas de los ejercicios 64-68, y luego calcule la inversa usando dicha fórmula.

69. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}$

70. La matriz del ejercicio 58

71. $\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$

72. $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 3 & 1 \\ -1 & 5 & 2 \end{bmatrix}$

3.4 La factorización *LU*

Así como es natural (e ilustrador) factorizar un número natural en un producto de otros números naturales (por ejemplo, $30 = 2 \cdot 3 \cdot 5$), frecuentemente también es útil factorizar matrices como productos de otras matrices. Cualquier representación de una matriz como producto de dos o más matrices se llama ***factorización de matrices***. Por ejemplo,

$$\begin{bmatrix} 3 & -1 \\ 9 & -5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 0 & -2 \end{bmatrix}$$

es una factorización de una matriz.

No es necesario decir que algunas factorizaciones son más útiles que otras. En esta sección se presenta una factorización matricial que surge en la solución de sistemas de ecuaciones lineales mediante eliminación gaussiana y es particularmente adecuada para su implementación en computadora. En capítulos posteriores se encontrarán otras factorizaciones matriciales igualmente útiles. De hecho, el tema es rico, y a él se han dedicado libros y cursos enteros.

Considere un sistema de ecuaciones lineales de la forma $A\mathbf{x} = \mathbf{b}$, donde A es una matriz de $n \times n$. La meta de este libro es demostrar que la eliminación gaussiana implícitamente factoriza A en un producto de matrices que permiten resolver fácilmente el sistema dado (y cualquier otro sistema con la misma matriz de coeficientes).

El siguiente ejemplo ilustra la idea básica.

Ejemplo 3.33

Sea

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$$

La reducción por renglones de A se realiza del modo siguiente:

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \xrightarrow[R_3 + R_1]{R_2 - 2R_1} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix} \xrightarrow{R_3 + 2R_2} \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = U \qquad (1)$$

Las tres matrices elementales E_1, E_2, E_3 que logran esta reducción de A a la forma escalonada U son (en orden):

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$

Por tanto,

$$E_3E_2E_1A = U$$

Al despejar para A, se obtiene

$$A = E_1^{-1}E_2^{-1}E_3^{-1}U = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}U$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}U = LU$$

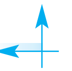
Por tanto, A puede factorizarse como

$$A = LU$$

donde U es una matriz *triangular superior* (vea los ejercicios para la sección 3.2) y L es *triangular inferior unitaria*. Esto es, L tiene la forma

$$L = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & 1 \end{bmatrix}$$

con ceros arriba y números 1 en la diagonal principal.



El ejemplo anterior motiva la siguiente definición.

El gran matemático inglés Alan M. Turing (1912–1954) introdujo la factorización *LU* en 1948, en un ensayo titulado "Rounding-off Errors en Matrix Processes" (*Quarterly Journal of Mechanics and Applied Mathematics, 1* (1948), pp. 287-308). Durante la segunda guerra mundial, Turing fue útil para descifrar el código "Enigma" alemán. Sin embargo, es mejor conocido por su trabajo en lógica matemática que tendió los cimientos teóricos para el desarrollo de las computadoras digitales y el moderno campo de la inteligencia artificial. La "prueba de Turing" que propuso en 1950 todavía se usa como uno de los hitos para abordar la cuestión de si una computadora puede considerarse "inteligente".

Definición Sea A una matriz cuadrada. Una factorización de A como $A = LU$, donde L es triangular inferior unitaria y U es triangular superior, se llama ***factorización LU*** de A.

Comentarios

- Observe que la matriz A en el ejemplo 3.33 tiene una factorización *LU* porque en la reducción por renglones de A *no se necesitaron intercambios de renglón*. Por tanto, todas las matrices elementales que surgieron fueron triangulares inferiores unitarias. Entonces, se garantiza que L es triangular inferior unitaria porque los inversos y productos

de matrices triangulares inferiores unitarias también son triangulares inferiores unitarias. (Vea los ejercicios 29 y 30.)

Si apareciera un cero en una posición pivote en algún paso, habría tenido que intercambiar renglones para obtener un pivote distinto de cero. Esto habría resultado en que



L ya no fuese triangular inferior unitaria. Más adelante se comentará más acerca de esta observación. (¿Puede encontrar una matriz para la cual será necesario el intercambio de renglones?)

- La noción de una factorización LU puede generalizarse a matrices no cuadradas al simplemente requerir que U sea una matriz en forma escalonada por renglones. (Vea los ejercicios 13 y 14.)
- Algunos libros definen una factorización LU de una matriz cuadrada A como cualquier factorización $A = LU$, donde L es triangular inferior y U es triangular superior.

El primer comentario anterior en esencia es una demostración del siguiente teorema.

Teorema 3.15 Si A es una matriz cuadrada que puede reducirse a forma escalonada sin usar intercambios de renglón, entonces A tiene una factorización LU.

Para ver por qué es útil la factorización LU, considere un sistema lineal $A\mathbf{x} = \mathbf{b}$, donde la matriz de coeficientes tiene una factorización LU, $A = LU$. El sistema $A\mathbf{x} = \mathbf{b}$ puede reescribirse como $LU\mathbf{x} = \mathbf{b}$ o $L(U\mathbf{x}) = \mathbf{b}$. Si ahora se define $\mathbf{y} = U\mathbf{x}$, entonces es posible resolver para $\mathbf{x}$ en dos etapas:

1. Resolver $L\mathbf{y} = \mathbf{b}$ para $\mathbf{y}$ mediante *sustitución hacia adelante* (vea los ejercicios 25 y 26 en la sección 2.1).
2. Resolver $U\mathbf{x} = \mathbf{y}$ para $\mathbf{x}$ mediante *sustitución hacia atrás*.

Cada uno de dichos sistemas lineales tiene solución directa porque las matrices coeficiente L y U son ambas triangulares. El siguiente ejemplo ilustra el método.

Ejemplo 3.34 Use una factorización LU de $A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix}$ para resolver $A\mathbf{x} = \mathbf{b}$, donde $\mathbf{b} = \begin{bmatrix} 1 \\ -4 \\ 9 \end{bmatrix}$.

Solución En el ejemplo 3.33, se encontró que

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = LU$$

Como se destaca líneas arriba, para resolver $A\mathbf{x} = \mathbf{b}$ (que es lo mismo que $L(U\mathbf{x}) = \mathbf{b}$), primero se resuelve $L\mathbf{y} = \mathbf{b}$ para $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$. Este es justo el sistema lineal

$$\begin{aligned} y_1 \qquad\qquad &= 1 \\ 2y_1 + y_2 \qquad &= -4 \\ -y_1 - 2y_2 + y_3 &= 9 \end{aligned}$$

La sustitución hacia adelante (esto es, al trabajar de arriba abajo) produce

$$y_1 = 1, y_2 = -4 - 2y_1 = -6, y_3 = 9 + y_1 + 2y_2 = -2$$

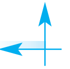
Por tanto, $\mathbf{y} = \begin{bmatrix} 1 \\ -6 \\ -2 \end{bmatrix}$ y ahora se resuelve $U\mathbf{x} = \mathbf{y}$ para $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$. Este sistema lineal es

$$\begin{aligned} 2x_1 + \quad x_2 + 3x_3 &= 1 \\ -3x_2 - 3x_3 &= -6 \\ 2x_3 &= -2 \end{aligned}$$

y la sustitución hacia atrás produce rápidamente

$$\begin{aligned} x_3 &= -1, \\ -3x_2 &= -6 + 3x_3 = -9 \text{ de manera que } x_2 = 3, \text{ y} \\ 2x_1 &= 1 - x_2 - 3x_3 = 1 \text{ de manera que } x_1 = \tfrac{1}{2} \end{aligned}$$

Por tanto, la solución al sistema dado $A\mathbf{x} = \mathbf{b}$ es $\mathbf{x} = \begin{bmatrix} \frac{1}{2} \\ 3 \\ -1 \end{bmatrix}$.



Una forma sencilla de encontrar factorizaciones *LU*

En el ejemplo 3.33 se calculó la matriz L como un producto de matrices elementales. Por fortuna, L puede calcularse directamente a partir del proceso de reducción por renglones sin la necesidad de calcular matrices elementales. Recuerde que se supone que A puede reducirse a forma escalonada por renglones sin usar intercambios de renglón. Si este es el caso, entonces todo el proceso de reducción por renglones puede hacerse usando solamente operaciones elementales con renglones de la forma $R_i - kR_j$. (¿Por qué no es necesario usar la operación elemental con los renglones restantes, y multiplicar un renglón por un escalar distinto de cero?) En la operación $R_i - kR_j$, el escalar k se denominará el ***multiplicador***.

En el ejemplo 3.33, las operaciones elementales con renglones que se usaron fueron, en orden,

$$\begin{aligned} &R_2 - 2R_1 && \text{(multiplicador} = 2) \\ &R_3 + R_1 = R_3 - (-1)R_1 && \text{(multiplicador} = -1) \\ &R_3 + 2R_2 = R_3 - (-2)R_2 && \text{(multiplicador} = -2) \end{aligned}$$

¡Los multiplicadores son precisamente las entradas de L que están bajo su diagonal! De hecho,

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}$$

y $L_{21} = 2$, $L_{31} = -1$ y $L_{32} = -2$. Note que la operación elemental con renglones $R_i - kR_j$ tiene su multiplicador k en la entrada (i, j) de L.

Ejemplo 3.35

Encuentre una factorización LU de

$$A = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix}$$

Solución Al reducir A a forma escalonada por renglones se obtiene

$$A = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix} \xrightarrow[R_4-(-3)R_1]{\substack{R_2-2R_1 \\ R_3-R_1}} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 1 & 2 & 3 \\ 0 & 8 & 7 & -16 \end{bmatrix}$$

$$\xrightarrow[R_4-4R_2]{R_3-\frac{1}{2}R_2} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & -1 & -8 \end{bmatrix}$$

$$\xrightarrow{R_4-(-1)R_3} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -4 \end{bmatrix} = U$$

Los primeros tres multiplicadores son 2, 1 y -3, y van en las entradas subdiagonales de la primera columna de L. De modo que, hasta el momento,

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & * & 1 & 0 \\ -3 & * & * & 1 \end{bmatrix}$$

Los siguientes dos multiplicadores son $\frac{1}{2}$ y 4, así que se sigue llenando L:

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & * & 1 \end{bmatrix}$$

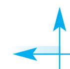
El multiplicador final, -1, sustituye al último * en L para producir

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & -1 & 1 \end{bmatrix}$$

Por tanto, una factorización LU de A es

$$A = \begin{bmatrix} 3 & 1 & 3 & -4 \\ 6 & 4 & 8 & -10 \\ 3 & 2 & 5 & -1 \\ -9 & 5 & -2 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 1 & \frac{1}{2} & 1 & 0 \\ -3 & 4 & -1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 3 & -4 \\ 0 & 2 & 2 & -2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -4 \end{bmatrix} = LU$$

como se comprueba fácilmente.



Comentarios

- Al aplicar este método es importante notar que las operaciones elementales con renglones $R_i - kR_j$ deben realizarse de arriba abajo dentro de cada columna (y usar la entrada diagonal como pivote) y columna por columna de izquierda a derecha. Para ilustrar lo que puede salir mal si no se obedecen estas reglas, considere la siguiente reducción por renglones:

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 1 \\ 2 & 2 & 1 \end{bmatrix} \xrightarrow{R_3 - 2R_2} \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & -1 \end{bmatrix} \xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & 2 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & -1 \end{bmatrix} = U$$

Esta vez el multiplicador se colocaría en L del modo siguiente: $L_{32} = 2$, $L_{21} = 1$. Se obtendría

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix}$$



pero $A \neq LU$. (¡Compruebe esto! Encuentre una factorización LU correcta de A.)

- Una forma alternativa de construir L es observar que los multiplicadores pueden obtenerse directamente de las matrices obtenidas en los pasos intermedios del proceso de reducción por renglones. En el ejemplo 3.33, examine los pivotes y las correspondientes columnas de las matrices que surgen en la reducción por renglones

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} \to A_1 = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 6 & 8 \end{bmatrix} \to \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = U$$

El primer pivote es 2, que aparece en la primera columna de A. Al dividir por el pivote las entradas de este vector columna que están en la diagonal o debajo de ella , se obtiene

$$\frac{1}{2}\begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

El siguiente pivote es -3, que aparece en la segunda columna de A_1. Al dividir por el pivote las entradas de este vector columna que están en la diagonal o debajo de ella, se obtiene

$$\frac{1}{(-3)}\begin{bmatrix} \\ -3 \\ 6 \end{bmatrix} = \begin{bmatrix} \\ 1 \\ -2 \end{bmatrix}$$

El pivote final (que no es necesario usar) es 2, en la tercera columna de U. Al dividir por el pivote las entradas de este vector columna que están en la diagonal o debajo de ella, se obtiene

$$\frac{1}{2}\begin{bmatrix} \\ \\ 2 \end{bmatrix} = \begin{bmatrix} \\ \\ 1 \end{bmatrix}$$

Si los tres vectores columna resultantes se colocan lado a lado en una matriz, se tiene

$$\begin{bmatrix} 1 & & \\ 2 & 1 & \\ -1 & -2 & 1 \end{bmatrix}$$

que es exactamente L una vez que las entradas sobre la diagonal se llenan con ceros.

En el capítulo 2 se remarcó que la forma escalonada por renglones de una matriz no es única. Sin embargo, si una matriz *invertible* A tiene una factorización LU, $A = LU$, entonces esta factorización es única.

Teorema 3.16 Si A es una matriz invertible que tiene una factorización LU, entonces L y U son únicas.

Demostración Suponga $A = LU$ y $A = L_1U_1$ son dos factorizaciones LU de A. Entonces $LU = L_1U_1$, donde L y L_1 son triangulares inferiores unitarias y U y U_1 son triangulares superiores. De hecho, U y U_1 son dos formas escalonadas por renglones (posiblemente diferentes) de A.

Por el ejercicio 30, L_1 es invertible. Dado que A es invertible, su forma escalonada reducida por renglones es una matriz identidad I por el teorema fundamental de matrices invertibles. Por tanto, U también se reduce por renglones a I (¿por qué?) y en consecuencia U también es invertible. Por lo cual,

$$L_1^{-1}(LU)U^{-1} = L_1^{-1}(L_1U_1)U^{-1} \quad \text{así} \quad (L_1^{-1}L)(UU^{-1}) = (L_1^{-1}L_1)(U_1U^{-1})$$

En consecuencia,

$$(L_1^{-1}L)I = I(U_1U^{-1}) \quad \text{entonces} \quad L_1^{-1}L = U_1U^{-1}$$

Pero $L_1^{-1}L$ es triangular inferior unitaria por el ejercicio 29 y U_1U^{-1} es triangular superior. (¿Por qué?) Se tiene que $L_1^{-1}L = U_1U^{-1}$ es *tanto* triangular inferior unitaria *como* triangular superior. La única matriz de esta forma es la matriz identidad, de modo que $L_1^{-1}L = I$ y $U_1U^{-1} = I$. Se tiene que $L = L_1$ y $U = U_1$, de modo que la factorización LU de A es única.

La factorización P^TLU

Ahora se explorará el problema de adaptar la factorización LU para manejar casos donde son necesarios los intercambios de renglón durante la eliminación gaussiana. Considere la matriz

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 3 & 6 & 2 \\ -1 & 1 & 4 \end{bmatrix}$$

Una reducción por renglón directa produce

$$A \to B = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 5 \\ 0 & 3 & 3 \end{bmatrix}$$

que no es una matriz triangular superior. Sin embargo, esto se puede convertir fácilmente en forma triangular superior al intercambiar los renglones 2 y 3 de B para obtener

$$U = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 3 & 3 \\ 0 & 0 & 5 \end{bmatrix}$$

Alternativamente, primero puede intercambiar los renglones 2 y 3 de A. Para este fin, sea P la matriz elemental

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

que corresponde a intercambiar los renglones 2 y 3, y sea E el producto de las matrices elementales que entonces reduce PA a U (de modo que $E^{-1} = L$ es triangular inferior unitaria). Por tanto, $EPA = U$, de modo que $A = (EP)^{-1}U = P^{-1}E^{-1}U = P^{-1}LU$.

Ahora esto sólo maneja el caso de *un solo* intercambio de renglón. En general, P será el producto $P = P_k \cdots P_2P_1$ de todas las matrices que intercambian renglón $P_1, P_2, \ldots, P_k$ (donde P_1 se realiza primero, etcétera). Tal matriz P se llama ***matriz permutación***. Observe que una matriz permutación surge de permutar los renglones de una matriz identidad en cierto orden. Por ejemplo, las siguientes son todas matrices permutación:

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Por fortuna, el inverso de una matriz permutación es fácil de calcular; de hecho, ¡no se necesita cálculo alguno!

Teorema 3.17 Si P es una matriz permutación, entonces $P^{-1} = P^T$.

Demostración Debe demostrar que $P^TP = I$. Pero el i-ésimo renglón de P^T es igual que la i-ésima columna de P, y ambos son iguales al mismo vector unitario estándar $\mathbf{e}$, porque P es una matriz permutación. De este modo

$$(P^TP)_{ii} = (i\text{-ésimo renglón de } P^T)(i\text{-ésima columna de } P) = \mathbf{e}^T\mathbf{e} = \mathbf{e} \cdot \mathbf{e} = 1$$

Esto muestra que las entradas diagonales de P^TP son todas 1. Por otra parte, si $j \neq i$, entonces la j-ésima columna de P es un vector unitario estándar *diferente* de $\mathbf{e}$, por decir $\mathbf{e}'$. Por tanto, una entrada típica fuera de la diagonal de P^TP está dado por

$$(P^TP)_{ij} = (i\text{-ésimo renglón de } P^T)(j\text{-ésima columna de } P) = \mathbf{e}^T\mathbf{e}' = \mathbf{e} \cdot \mathbf{e}' = 0$$

En consecuencia, P^TP es una matriz identidad, como se quería demostrar.

Por tanto, en general, puede factorizar una matriz cuadrada A como $A = P^{-1}LU = P^TLU$.

Definición Sea A una matriz cuadrada. Una factorización de A como $A = P^TLU$, donde P es una matriz permutación, L es triangular inferior unitaria y U es triangular superior, se llama ***factorización P^TLU*** de A.

Ejemplo 3.36

Encuentre una factorización P^TLU de $A = \begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}$.

Solución Primero reduzca A a forma escalonada por renglones. Claramente, es necesario al menos un intercambio de renglón.

$$A = \begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 6 \\ 2 & 1 & 4 \end{bmatrix} \xrightarrow{R_3 - 2R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 6 \\ 0 & -3 & -2 \end{bmatrix}$$

$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & 6 \end{bmatrix}$$

Se usaron dos intercambios de renglón ($R_1 \leftrightarrow R_2$ y luego $R_2 \leftrightarrow R_3$), de modo que la matriz permutación requerida es

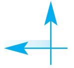
$$P = P_2P_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

Ahora se encuentra una factorización LU de PA.

$$PA = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 6 \\ 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \\ 0 & 0 & 6 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & 6 \end{bmatrix} = U$$

Por tanto, $L_{21} = 2$, y por tanto

$$A = P^TLU = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -2 \\ 0 & 0 & 6 \end{bmatrix}$$



La discusión anterior justifica el siguiente teorema.

Teorema 3.18 Toda matriz cuadrada tiene una factorización $P^T LU$.

Comentario Incluso para una matriz invertible, la factorización $P^T LU$ no es única. En el ejemplo 3.36, un solo intercambio de renglón $R_1 \leftrightarrow R_3$ también habría funcionado, lo que conduciría a una P diferente. Sin embargo, una vez determinada P, L y U son únicas.

Consideraciones de cálculo

Si A es de $n \times n$, entonces el número total de operaciones (multiplicaciones y divisiones) requeridas para resolver un sistema lineal $A\mathbf{x} = \mathbf{b}$ usando una factorización LU de A, es $T(n) \approx n^3/3$, el mismo que se requiere para la eliminación gaussiana. (Vea la Exploración "Operaciones de conteo", en el capítulo 2.) Esto difícilmente es sorprendente, pues la fase de eliminación hacia adelante produce la factorización LU en $\approx n^3/3$ pasos, mientras que las sustituciones hacia adelante y hacia atrás requieren $\approx n^2/2$ pasos. Por tanto, para valores grandes de n, el término $n^3/3$ es dominante. Desde este punto de vista, entonces, la eliminación gaussiana y la factorización LU son equivalentes.

Sin embargo, la factorización LU tiene otras ventajas:

- Desde un punto de vista de almacenamiento, la factorización LU es muy compacta, porque puede *sobreescribir* las entradas de A con las entradas de L y U *conforme se calculan*. En el ejemplo 3.33 se encontró que

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 4 & -1 & 3 \\ -2 & 5 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & -3 \\ 0 & 0 & 2 \end{bmatrix} = LU$$

Esto puede almacenarse como

$$\begin{bmatrix} 2 & -1 & 3 \\ 2 & -3 & -3 \\ -1 & -2 & 2 \end{bmatrix}$$

con las entradas colocadas en el oren $(1, 1), (1, 2), (1, 3), (2, 1), (3, 1), (2, 2), (2, 3), (3, 2), (3, 3)$. En otras palabras, las entradas subdiagonales de A se sustituyen con los multiplicadores correspondientes. (¡Compruebe que esto funciona!)

- Una vez calculada la factorización A de LU, puede usarse para resolver tantos sistemas lineales de la forma $A\mathbf{x} = \mathbf{b}$ como se quiera. Sólo necesita aplicar el método del ejemplo 3.34 y variar el vector $\mathbf{b}$ cada vez.
- Para matrices con ciertas formas especiales, sobre todo aquellas con un gran número de ceros (las llamadas matrices "dispersas") concentrados fuera de la diagonal, existen métodos que simplificarán el cálculo de una factorización *LU*. En estos casos, este método es más rápido que la eliminación gaussiana para resolver $A\mathbf{x} = \mathbf{b}$.
- Para una matriz invertible A, puede usarse una factorización *LU* de A para encontrar A^{-1}, si es necesario. Más aún, esto puede hacerse en tal forma que simultáneamente produzca una factorización de A^{-1}. (Vea los ejercicios 15-18.)

Comentario Si tiene un CAS (como MATLAB) que tenga incluida la factorización *LU*, puede notar algunas diferencias entre sus cálculos a mano y el resultado de la computadora. Esto se debe a que la mayoría de los CAS automáticamente tratarán de realizar pivoteo parcial para reducir errores de redondeo. (Vea la exploración "Pivoteo parcial" en el capítulo 2.) El ensayo de Turing es una discusión extensa de tales errores en el contexto de las factorizaciones de matrices.

Esta sección sirvió para presentar una de las factorizaciones matriciales más útiles. En capítulos posteriores se encontrarán otras igualmente útiles.

Ejercicios 3.4

En los ejercicios 1-6, resuelva el sistema $A\mathbf{x} = \mathbf{b}$ *usando la factorización LU dada de A.*

1. $A = \begin{bmatrix} -2 & 1 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -2 & 1 \\ 0 & 6 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

2. $A = \begin{bmatrix} 4 & -2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix}\begin{bmatrix} 4 & -2 \\ 0 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 8 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & 1 & -2 \\ -2 & 3 & -4 \\ 4 & -3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -\frac{5}{4} & 1 \end{bmatrix}$
$\times \begin{bmatrix} 2 & 1 & -2 \\ 0 & 4 & -6 \\ 0 & 0 & -\frac{7}{2} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -3 \\ 1 \\ 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 2 & -4 & 0 \\ 3 & -1 & 4 \\ -1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ -\frac{1}{2} & 0 & 1 \end{bmatrix}$
$\times \begin{bmatrix} 2 & -4 & 0 \\ 0 & 5 & 4 \\ 0 & 0 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ -5 \end{bmatrix}$

5. $A = \begin{bmatrix} 2 & -1 & 0 & 0 \\ 6 & -4 & 5 & -3 \\ 8 & -4 & 1 & 0 \\ 4 & -1 & 0 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ 4 & 0 & 1 & 0 \\ 2 & -1 & 5 & 1 \end{bmatrix}$
$\times \begin{bmatrix} 2 & -1 & 0 & 0 \\ 0 & -1 & 5 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 4 & 3 & 0 \\ -2 & -5 & -1 & 2 \\ 3 & 6 & -3 & -4 \\ -5 & -8 & 9 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ 3 & -2 & 1 & 0 \\ -5 & 4 & -2 & 1 \end{bmatrix}$
$\times \begin{bmatrix} 1 & 4 & 3 & 0 \\ 0 & 3 & 5 & 2 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ -3 \\ -1 \\ 0 \end{bmatrix}$

En los ejercicios 7-12, encuentre una factorización LU de la matriz dada.

7. $\begin{bmatrix} 1 & 2 \\ -3 & -1 \end{bmatrix}$

8. $\begin{bmatrix} 2 & -4 \\ 3 & 1 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 8 & 7 & 9 \end{bmatrix}$ **10.** $\begin{bmatrix} 2 & 2 & -1 \\ 4 & 0 & 4 \\ 3 & 4 & 4 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 2 & 3 & -1 \\ 2 & 6 & 3 & 0 \\ 0 & 6 & -6 & 7 \\ -1 & -2 & -9 & 0 \end{bmatrix}$

12. $\begin{bmatrix} 2 & 2 & 2 & 1 \\ -2 & 4 & -1 & 2 \\ 4 & 4 & 7 & 3 \\ 6 & 9 & 5 & 8 \end{bmatrix}$

Generalice la definición de factorización LU a matrices no cuadradas al simplemente requerir que U sea una matriz en forma escalonada por renglones. Con esta modificación, encuentre una factorización LU de las matrices en los ejercicios 13 y 14.

13. $\begin{bmatrix} 1 & 0 & 1 & -2 \\ 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 2 & 0 & -1 & 1 \\ -2 & -7 & 3 & 8 & -2 \\ 1 & 1 & 3 & 5 & 2 \\ 0 & 3 & -3 & -6 & 0 \end{bmatrix}$

Para una matriz invertible con una factorización LU, $A = LU$, L y U serán invertibles y $A^{-1} = U^{-1}L^{-1}$. En los ejercicios 15 y 16, encuentre L^{-1}, U^{-1} y A^{-1} para la matriz dada.

15. A en el ejercicio 1 **16.** A en el ejercicio 4

El inverso de una matriz también se puede calcular al resolver varios sistemas de ecuaciones usando el método del ejemplo 3.34. Para una matriz A de $n \times n$, para encontrar su inverso es necesario resolver $AX = I_n$ para la matriz X de $n \times n$. Al escribir esta ecuación como $A[\mathbf{x}_1 \; \mathbf{x}_2 \cdots \mathbf{x}_n] = [\mathbf{e}_1 \; \mathbf{e}_2 \cdots \mathbf{e}_n]$, y usar la forma matriz-columna de AX, se ve que es necesario resolver n sistemas de ecuaciones lineales. $A\mathbf{x}_1 = \mathbf{e}_1, A\mathbf{x}_2 = \mathbf{e}_2, \ldots, A\mathbf{x}_n = \mathbf{e}_n$. Más aún, puede usar la factorización $A = LU$ para resolver cada uno de dichos sistemas.

En los ejercicios 17 y 18, use el enfoque recién destacado para encontrar A^{-1} para la matriz dada. Compare con el método de los ejercicios 15 y 16.

17. A en el ejercicio 1 **18.** A en el ejercicio 4

En los ejercicios 19-22, escriba la matriz permutación dada como un producto de matrices elementales (intercambio de renglón).

19. $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ **20.** $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

21. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ **22.** $\begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 23-25, encuentre una factorización P^TLU de la matriz A dada.

23. $A = \begin{bmatrix} 0 & 1 & 4 \\ -1 & 2 & 1 \\ 1 & 3 & 3 \end{bmatrix}$ **24.** $A = \begin{bmatrix} 0 & 0 & 1 & 2 \\ -1 & 1 & 3 & 2 \\ 0 & 2 & 1 & 1 \\ 1 & 1 & -1 & 0 \end{bmatrix}$

25. $A = \begin{bmatrix} 0 & -1 & 1 & 3 \\ -1 & 1 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

26. Demuestre que existen exactamente $n!$ matrices permutación de $n \times n$.

En los ejercicios 27-28, resuelva el sistema $A\mathbf{x} = \mathbf{b}$ usando la factorización dada $A = P^TLU$. Puesto que $PP^T = I$, $P^TLU\mathbf{x} = \mathbf{b}$ puede reescribirse como $LU\mathbf{x} = P\mathbf{b}$. Entonces este sistema puede resolverse con el método del ejemplo 3.34.

27. $A = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 3 & 2 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$

$\times \begin{bmatrix} 2 & 3 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & -\frac{5}{2} \end{bmatrix} = P^TLU, \; \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 5 \end{bmatrix}$

28. $A = \begin{bmatrix} 8 & 3 & 5 \\ 4 & 1 & 2 \\ 4 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}$

$\times \begin{bmatrix} 4 & 1 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix} = P^TLU, \; \mathbf{b} = \begin{bmatrix} 16 \\ -4 \\ 4 \end{bmatrix}$

29. Demuestre que un producto de matrices triangulares inferiores unitarias es triangular inferior unitaria.

30. Demuestre que toda matriz triangular inferior unitaria es invertible y que su inversa también es triangular inferior unitaria.

*Una **factorización LDU** de una matriz cuadrada A es una factorización A = LDU, donde L es una matriz triangular inferior unitaria, D es una matriz diagonal y U es una matriz triangular superior unitaria (triangular superior con 1 en su diagonal). En los ejercicios 31 y 32, encuentre una factorización LDU de A.*

31. A en el ejercicio 1 **32.** A en el ejercicio 4

33. Si A es simétrica e invertible y tiene una factorización LDU, demuestre que $U = L^T$.

34. Si A es simétrica e invertible, y $A = LDL^T$ (con L triangular inferior unitaria y D diagonal), demuestre que esta factorización es única. Esto es: demuestre que, si también se tiene $A = L_1D_1L_1^T$ (con L_1 triangular inferior unitaria y D_1 diagonal), entonces $L = L_1$ y $D = D_1$.

3.5 Subespacios, bases, dimensión y rank

Esta sección presenta tal vez las ideas más importantes de todo el libro. Ya se vio que existe una interacción entre geometría y álgebra: con frecuencia es posible usar intuición geométrica y razonar para obtener resultados algebraicos, y la potencia del álgebra con frecuencia permitirá extender los hallazgos más allá de los escenarios geométricos en los que surgieron por primera vez.

En el estudio de los vectores ya encontró informalmente todos los conceptos en esta sección. Aquí, comenzará a volverse más formal al proporcionar definiciones para las ideas clave. Como verá, la noción de *subespacio* es simplemente una generalización algebraica de los ejemplos geométricos de rectas y planos que pasan por el origen. El concepto fundamental de *base* para un subespacio se deriva entonces de la idea de vectores directores para tales rectas y planos. El concepto de base le permitirá dar una definición precisa de *dimensión* que concuerde con una idea geométrica intuitiva del término, aunque es suficientemente flexible para permitir la generalización a otros escenarios.

También comenzará a ver que dichas ideas arrojan más luz a lo que ya sabe acerca de las matrices y la solución de sistemas de ecuaciones lineales. En el capítulo 6 encontrará nuevamente todas estas ideas fundamentales con más detalle. Considere esta sección como una sesión de "familiarización".

Un plano que pasa por el origen en $\mathbb{R}^3$ "parece" una copia de $\mathbb{R}^2$. Intuitivamente, estaría de acuerdo en que ambos son "bidimensionales". Si presiona un poco, también puede decir que cualquier cálculo que pueda hacerse con vectores en $\mathbb{R}^2$ también puede realizarse en un plano que pasa por el origen. En particular, puede sumar y tomar múltiplos escalares (y, más generalmente, formar combinaciones lineales) de vectores en tal plano, y los resultados son otros vectores *en el mismo plano*. Se dice que, como $\mathbb{R}^2$, un plano que pasa por el origen es *cerrado* con respecto a las operaciones de suma y multiplicación escalar. (Vea la figura 3.2.)

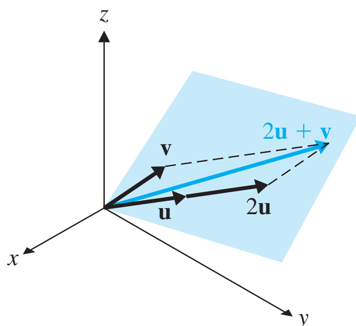


Figura 3.2

Pero los vectores en este plano, ¿son objetos bidimensionales o tridimensionales? Puede argumentar que son tridimensionales porque existen en $\mathbb{R}^3$ y por tanto tienen tres componentes. Por otra parte, pueden describirse como una combinación lineal de sólo dos vectores, vectores dirección para el plano, y por tanto son objetos bidimensionales que existen en un plano bidimensional. La noción de subespacio es la clave para resolver este rompecabezas.

Definición Un ***subespacio*** de $\mathbb{R}^n$ es cualquier colección S de vectores en $\mathbb{R}^n$ tal que:

1. El vector cero $\mathbf{0}$ está en S.
2. Si $\mathbf{u}$ y $\mathbf{v}$ están en S, entonces $\mathbf{u} + \mathbf{v}$ está en S. (S es ***cerrado bajo la suma***.)
3. Si $\mathbf{u}$ está en S y c es un escalar, entonces $c\mathbf{u}$ está en S. (S es ***cerrado bajo la multiplicación por un escalar***.)

También podría combinar las propiedades (2) y (3) y requerir, de manera equivalente, que S sea ***cerrado bajo combinaciones lineales***:

$$\text{Si } \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_k \text{ están en } S \text{ y } c_1, c_2, \ldots, c_k \text{ son escalares,}$$
$$\text{entonces } c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_k\mathbf{u}_k \text{ están en } S.$$

Ejemplo 3.37

Toda recta y plano a través del origen en $\mathbb{R}^3$ es un subespacio de $\mathbb{R}^3$. Debe ser claro geométricamente que se satisfacen las propiedades (1) a (3). He aquí una denostración algebraica en el caso de un plano a través del origen. En el ejercicio 9 se le pide dar la demostración correspondiente para una recta.

Sea $\mathcal{P}$ un plano a través del origen con vectores directores $\mathbf{v}_1$ y $\mathbf{v}_2$. Por tanto, $\mathcal{P} = \text{gen}(\mathbf{v}_1, \mathbf{v}_2)$. El vector cero $\mathbf{0}$ está en $\mathcal{P}$, pues $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2$. Ahora sean

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \quad \text{y} \quad \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2$$

dos vectores en $\mathcal{P}$. Entonces

$$\mathbf{u} + \mathbf{v} = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2) + (d_1\mathbf{v}_1 + d_2\mathbf{v}_2) = (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2$$

Por tanto, $\mathbf{u} + \mathbf{v}$ es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$,y por tanto está en $\mathcal{P}$.

Ahora sea c un escalar. Entonces

$$c\mathbf{u} = c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2$$

que muestra que $c\mathbf{u}$ también es una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$ y por tanto está en $\mathcal{P}$. Se demostró que $\mathcal{P}$ satisface las propiedades (1) a (3) y en consecuencia es un subespacio de $\mathbb{R}^3$.

Si observa cuidadosamente los detalles del ejemplo 3.37, notará que el hecho de que $\mathbf{v}_1$ y $\mathbf{v}_2$ fueran vectores en $\mathbb{R}^3$ no tuvo papel alguno en la verificación de las propiedades. En consecuencia, el método algebraico utilizado debe generalizarse más allá de $\mathbb{R}^3$ y aplicarse en situaciones donde ya no se pueda visualizar la geometría. Lo hace. Más aún, el método del ejemplo 3.37 puede servir como "plantilla" en escenarios más generales. Cuando el ejemplo 3.37 se generaliza al generador de un conjunto arbitrario de vectores en cualquier $\mathbb{R}^n$, el resultado es suficientemente importante como para llamarse teorema.

Teorema 3.19 Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ vectores en $\mathbb{R}^n$. Entonces $\text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ es un subespacio de $\mathbb{R}^n$.

Demostración Sea $S = \text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$. Para demostrar la propiedad (1) de la definición, simplemente se observa que el vector cero $\mathbf{0}$ está en S, ya que $\mathbf{0} = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k$.

Ahora sean

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \quad \text{y} \quad \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k$$

dos vectores en S. Entonces

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) + (d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k)\\ &= (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2 + \cdots + (c_k + d_k)\mathbf{v}_k\end{aligned}$$

Por tanto, $\mathbf{u} + \mathbf{v}$ es una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y por tanto está en S. Esto verifica la propiedad (2).

Para demostrar la propiedad (3), sea c un escalar. Entonces

$$\begin{aligned}c\mathbf{u} &= c(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k)\\ &= (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2 + \cdots + (cc_k)\mathbf{v}_k\end{aligned}$$

lo que demuestra que $c\mathbf{u}$ también es una combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y por tanto está en S. Se demostró que S satisface las propiedades (1) a (3) y en consecuencia es un subespacio de $\mathbb{R}^n$.

A gen$(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ se le llamará ***subespacio generado por*** $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Con frecuencia podrá ahorrar mucho trabajo al reconocer cuándo puede aplicarse el Teorema 3.19.

Ejemplo 3.38

Demuestre que el conjunto de todos los vectores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que satisfacen las condiciones $x = 3y$ y $z = -2y$ forman un subespacio de $\mathbb{R}^3$.

Solución La sustitución de las dos condiciones en $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ produce

$$\begin{bmatrix} 3y \\ y \\ -2y \end{bmatrix} = y\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$$

Dado que y es arbitrario, el conjunto dado de vectores es gen$\left(\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}\right)$ y por ende es un subespacio de $\mathbb{R}^3$, por el Teorema 3.19.

Geométricamente, el conjunto de vectores del ejemplo 3.38 representa la recta que pasa por el origen en $\mathbb{R}^3$, con vector director $\begin{bmatrix} 3 \\ 1 \\ -2 \end{bmatrix}$.

Ejemplo 3.39

Determine si el conjunto de todos los vectores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ que satisface las condiciones $x = 3y + 1$ y $z = -2y$ es un subespacio de $\mathbb{R}^3$.

Solución Esta vez se tienen todos los vectores de la forma

$$\begin{bmatrix} 3y + 1 \\ y \\ -2y \end{bmatrix}$$

El vector cero no es de esta forma. (¿Por qué no? Trate de resolver $\begin{bmatrix} 3y + 1 \\ y \\ -2y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.)

Por tanto, la propiedad (1) no se cumple, de modo que este conjunto no puede ser un subespacio de $\mathbb{R}^3$.

Ejemplo 3.40

Determine si el conjunto de todos los vectores $\begin{bmatrix} x \\ y \end{bmatrix}$, donde $y = x^2$, es un subespacio de $\mathbb{R}^2$.

Solución Se trata de los vectores de la forma $\begin{bmatrix} x \\ x^2 \end{bmatrix}$, llame S a este conjunto. Esta vez $\mathbf{0} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ pertenece a S (tome $x = 0$), de modo que se cumple la propiedad (1). Sea $\mathbf{u} = \begin{bmatrix} x_1 \\ x_1^2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} x_2 \\ x_2^2 \end{bmatrix}$ en S. Entonces

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} x_1 + x_2 \\ x_1^2 + x_2^2 \end{bmatrix}$$

que, en general, no está en S, pues no tiene la forma correcta; esto es: $x_1^2 + x_2^2 \neq (x_1 + x_2)^2$. Para ser específico, busque un contraejemplo. Si

$$\mathbf{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

entonces tanto **u** como **v** están en S, pero su suma $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$ no está en S pues $5 \neq 3^2$.

Por tanto, la propiedad (2) falla y S no es un subespacio de $\mathbb{R}^2$.

Comentario Para que un conjunto S sea un subespacio de algún $\mathbb{R}^n$, debe *demostrar* que las propiedades (1) a (3) se sustentan *en general*. Sin embargo, para que S *falle* en ser subespacio de $\mathbb{R}^n$, es suficiente demostrar que no se cumple *una* de las tres propiedades. La ruta más sencilla por lo general es encontrar un solo *contraejemplo* específico para ilustrar la falla de la propiedad. Una vez hecho esto, no hay necesidad de considerar las otras propiedades.

Subespacios asociados con matrices

En el contexto de las matrices surgen una gran cantidad de ejemplos de subespacios. Ya se encontraron las más importantes en el capítulo 2; ahora se revisarán nuevamente con la noción de subespacio en mente.

Definición Sea A una matriz de $m \times n$.

1. El ***espacio renglón*** de A es el subespacio renglón(A) de $\mathbb{R}^n$ generado por los renglones de A.
2. El ***espacio columna*** de A es el subespacio col(A) de $\mathbb{R}^m$ generado por las columnas de A.

Ejemplo 3.41 Considere la matriz

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 3 & -3 \end{bmatrix}$$

(a) Determine si $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ está en el espacio columna de A.

(b) Determine si $\mathbf{w} = [4 \quad 5]$ está en el espacio renglón de A.
(c) Describa renglón(A) y col(A).

Solución
(a) Por el Teorema 2.4 y el análisis que le precedió, $\mathbf{b}$ es una combinación lineal de las columnas de A si y sólo si el sistema lineal $A\mathbf{x} = \mathbf{b}$ es consistente. La matriz aumentada se reduce por renglones del modo siguiente:

$$\left[\begin{array}{cc|c} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 3 & -3 & 3 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{array}\right]$$

Por tanto, el sistema es consistente (y, de hecho, tiene una solución única). Por tanto, $\mathbf{b}$ está en col(A). (De hecho este ejemplo es el 2.18, parafraseado en la terminología de esta sección.)

(b) Como también se vio en la sección 2.3, las operaciones elementales con renglones simplemente crean combinaciones lineales de los renglones de una matriz. Esto es, producen vectores solamente en el espacio renglón de la matriz. Si el vector $\mathbf{w}$ está en renglón(A), entonces $\mathbf{w}$ es una combinación lineal de los renglones de A, de modo que si A se aumenta con $\mathbf{w}$ como $\left[\begin{array}{c} A \\ \hline \mathbf{w} \end{array}\right]$, será posible aplicar operaciones elementales con renglones a esta matriz aumentada para reducirla a la forma $\left[\begin{array}{c} A' \\ \hline \mathbf{0} \end{array}\right]$ usando solamente operaciones elementales con renglones de la forma $R_i + kR_j$, donde $i > j$; en otras palabras, *trabajar de arriba abajo en cada columna*. (¿Por qué?)

En este ejemplo se tiene

$$\left[\begin{array}{c} A \\ \hline \mathbf{w} \end{array}\right] = \left[\begin{array}{cc} 1 & -1 \\ 0 & 1 \\ 3 & -3 \\ \hline 4 & 5 \end{array}\right] \xrightarrow[R_4 - 4R_1]{R_3 - 3R_1} \left[\begin{array}{cc} 1 & -1 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 9 \end{array}\right] \xrightarrow{R_4 - 9R_2} \left[\begin{array}{cc} 1 & -1 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 0 \end{array}\right]$$

Por tanto, **w** es una combinación lineal de los renglones de A (de hecho, estos cálculos
 muestran que $\mathbf{w} = 4[1 \quad -1] + 9[0 \quad 1]$; ¿cómo?), y por tanto **w** está en renglón(A).

(c) Es fácil comprobar que, para cualquier vector $\mathbf{w} = [x \quad y]$, la matriz aumentada $\left[\begin{array}{c} A \\ \hline \mathbf{w} \end{array}\right]$ se reduce a

$$\left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ \hline 0 & 0 \end{array}\right]$$

en forma similar. Por tanto, todo vector en $\mathbb{R}^2$ está en renglón(A), y en consecuencia renglón $(A) = \mathbb{R}^2$.

Encontrar col(A) es idéntico a resolver el ejemplo 2.21, en que se determinó que coincide con el plano (a través del origen) en $\mathbb{R}^3$ con ecuación $3x - z = 0$. (Dentro de poco se descubrirán otras formas de responder este tipo de preguntas.)



Comentario También podría responder el inciso (b) y la primera parte del inciso (c) al observar que cualquier pregunta acerca de los *renglones* de A es la correspondiente pregunta acerca de las *columnas* de A^T. De este modo, por ejemplo **w** está en renglón(A) si y sólo si $\mathbf{w}^T$ está en col(A^T). Esto es verdadero si y sólo si el sistema $A^T\mathbf{x} = \mathbf{w}^T$ es consistente. Ahora puede proceder como en el inciso (a). (Vea los ejercicios 21-24.)

Los comentarios hechos acerca de la relación entre operaciones elementales con renglones y el espacio renglón se resumen en el siguiente teorema.

Teorema 3.20 Sea B cualquier matriz que es equivalente por renglones a una matriz A. Entonces renglón(B) = renglón(A).

Demostración La matriz A puede transformarse en B mediante una secuencia de operaciones con renglones. En consecuencia, los renglones de B son combinaciones lineales de los renglones de A; por tanto, las combinaciones lineales de los renglones de B son combinaciones lineales de los renglones de A. (Vea el ejercicio 21 en la sección 2.3.) Se tiene que renglón(B) $\subseteq$ renglón(A).

Por otra parte, invertir dichas operaciones con renglones transforman B en A. Por tanto, el argumento anterior demuestra que renglón(A) $\subseteq$ renglón(B). Combinando estos resultados renglón(A) = renglón(B)

Existe otro importante subespacio que ya se encontró: el conjunto de soluciones de un sistema homogéneo de ecuaciones lineales. Es fácil probar que este subespacio satisface las tres propiedades de los subespacios.

Teorema 3.21 Sea A una matriz de $m \times n$ y sea N el conjunto de soluciones del sistema lineal homogéneo $A\mathbf{x} = \mathbf{0}$. Entonces N es un subespacio de $\mathbb{R}^n$.

Demostración [Note que **x** debe ser un vector (columna) en $\mathbb{R}^n$ para que A**x** esté definido y que $\mathbf{0} = \mathbf{0}_m$ es el vector cero en $\mathbb{R}^m$.] Dado que $A\mathbf{0}_n = \mathbf{0}_m$, $\mathbf{0}_n$ está en N. Ahora sean **u** y **v** en N. Por tanto, $A\mathbf{u} = \mathbf{0}$ y $A\mathbf{v} = \mathbf{0}$. Se tiene que

$$A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

En consecuencia, $\mathbf{u} + \mathbf{v}$ está en N. Finalmente, para cualquier escalar c,

$$A(c\mathbf{u}) = c(A\mathbf{u}) = c\mathbf{0} = \mathbf{0}$$

y en consecuencia $c\mathbf{u}$ también está en N. Se tiene que N es un subespacio de $\mathbb{R}^n$.

Definición Sea A una matriz de $m \times n$. El ***espacio nulo*** de A es el subespacio de $\mathbb{R}^n$ que consiste de las soluciones del sistema lineal homogéneo $A\mathbf{x} = \mathbf{0}$. Se denota mediante nul(A).

El hecho de que el espacio nulo de una matriz sea un subespacio permite probar lo que la intuición y los ejemplos condujeron a entender acerca de las soluciones de los sistemas lineales: no tienen solución, tienen una solución única o un número infinito de soluciones.

Teorema 3.22 Sea A una matriz cuyas entradas son números reales. Para cualquier sistema de ecuaciones lineales $A\mathbf{x} = \mathbf{b}$, exactamente uno de los siguientes es verdadero:

a. No hay solución.
b. Hay una solución única.
c. Hay un número infinito de soluciones.

A primera vista, no es completamente claro cómo se debe proceder para probar este teorema. Un poco de reflexión debe persuadirlo de que lo que realmente se pide probar es que si (a) y (b) no son verdaderas, entonces (c) es la única posibilidad. Esto es, si hay más de una solución, entonces no puede haber sólo dos o incluso muchas en una cantidad finita, sino que debe haber un número infinito.

Demostración Si el sistema $A\mathbf{x} = \mathbf{b}$ no tiene solución o tiene exactamente una solución, está demostrado. Suponga, entonces, que existen al menos dos soluciones distintas de $A\mathbf{x} = \mathbf{b}$, por ejemplo $\mathbf{x}_1$ y $\mathbf{x}_2$. Por tanto,

$$A\mathbf{x}_1 = \mathbf{b} \quad \text{y} \quad A\mathbf{x}_2 = \mathbf{b}$$

con $\mathbf{x}_1 \neq \mathbf{x}_2$. Se tiene que

$$A(\mathbf{x}_1 - \mathbf{x}_2) = A\mathbf{x}_1 - A\mathbf{x}_2 = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

Sea $\mathbf{x}_0 = \mathbf{x}_1 - \mathbf{x}_2$. Entonces $\mathbf{x}_0 \neq \mathbf{0}$ y $A\mathbf{x}_0 = \mathbf{0}$. En consecuencia, el espacio nulo de A es no trivial, y dado que nul(A) es cerrado para la multiplicación por un escalar, $c\mathbf{x}_0$ está en nul(A) para todo escalar c. En consecuencia, el espacio nulo de A contiene infinitos vectores (dado que contiene *al menos* a todo vector de la forma $c\mathbf{x}_0$, y existen infinitos de ellos.)

Ahora, considere los vectores (un número infinito) de la forma $\mathbf{x}_1 + c\mathbf{x}_0$, cuando c varía a lo largo del conjunto de los números reales. Se tiene

$$A(\mathbf{x}_1 + c\mathbf{x}_0) = A\mathbf{x}_1 + cA\mathbf{x}_0 = \mathbf{b} + c\mathbf{0} = \mathbf{b}$$

Por tanto, existe un número infinito de soluciones de la ecuación $A\mathbf{x} = \mathbf{b}$.

Bases

Es posible extraer un poco más de la idea intuitiva de que los subespacios son generalizaciones de planos a través del origen en $\mathbb{R}^3$. Un plano se genera mediante cualesquiera

dos vectores que son paralelos al plano pero no son paralelos entre ellos. En lenguaje algebraico, dos de tales vectores generan el plano y son linealmente independientes. Menos de dos vectores no funcionarán; más de dos vectores no son necesarios. Esta es la esencia de una *base* para un subespacio.

Definición Una ***base*** para un subespacio S de $\mathbb{R}^n$ es un conjunto de vectores en S que

1. genera S y
2. es linealmente independiente.

Ejemplo 3.42

En la sección 2.3 se vio que los vectores unitarios estándar $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$ en $\mathbb{R}^n$ son linealmente independientes y generan $\mathbb{R}^n$. Por tanto, forman una base para $\mathbb{R}^n$, llamada ***base estándar***.

Ejemplo 3.43

En el ejemplo 2.19 se demostró que $\mathbb{R}^2 = \text{gen}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$. Dado que $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$ y $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ también son linealmente independientes (ya que no son múltiplos), forman una base para $\mathbb{R}^2$.

Un subespacio puede tener (y tendrá) más de una base. Por ejemplo, recién se vio que $\mathbb{R}^2$ tiene la base estándar $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ y la base $\left\{\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}\right\}$. Sin embargo, dentro de poco se probará que el *número* de vectores en una base para un subespacio dado siempre será el mismo.

Ejemplo 3.44

Encuentre una base para $S = \text{gen}(\mathbf{u}, \mathbf{v}, \mathbf{w})$, donde

$$\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, \; \mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \quad \text{y} \quad \mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix}$$

Solución Los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ ya generan S, de modo que serán una base de S si también son linealmente independientes. Es fácil determinar que no lo son; de hecho, $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$. Por tanto, puede ignorarse $\mathbf{w}$, pues cualquier combinación lineal que involucre $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ puede reescribirse para involucrar sólo a $\mathbf{u}$ y $\mathbf{v}$. (Vea también el ejercicio 47 de la sección 2.3.) Esto implica que $S = \text{gen}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \text{gen}(\mathbf{u}, \mathbf{v})$, y dado que $\mathbf{u}$ y $\mathbf{v}$ ciertamente son linealmente independientes (¿por qué?), forman una base para S. (Geométricamente, esto significa que $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ yacen todos en el mismo plano, y $\mathbf{u}$ y $\mathbf{v}$ pueden servir como un conjunto de vectores directores para este plano.)

Ejemplo 3.45

Encuentre una base para el espacio renglón de

$$A = \begin{bmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{bmatrix}$$

Solución La forma escalonada reducida por renglones de A es

$$R = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Por el Teorema 3.20, renglón(A) = renglón(R), de modo que es suficiente encontrar una base para el espacio renglón de R. Pero renglón(R) claramente es generado por sus renglones distintos de cero, y es fácil comprobar que el patrón en escalera fuerza a los primeros tres renglones de R a ser linealmente independientes. (Este es un hecho general, que necesitará establecer para probar el ejercicio 33.) En consecuencia, una base para el espacio renglón de A es

$$\{[1 \quad 0 \quad 1 \quad 0 \quad -1], [0 \quad 1 \quad 2 \quad 0 \quad 3], [0 \quad 0 \quad 0 \quad 1 \quad 4]\}$$

Puede usar el método del ejemplo 3.45 para encontrar una base para el subespacio generado por un conjunto dado de vectores.

Ejemplo 3.46

Vuelva a trabajar el ejemplo 3.44 usando el método del ejemplo 3.45.

Solución Transponga $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ para obtener vectores renglón y luego formar una matriz con dichos vectores como sus renglones:

$$B = \begin{bmatrix} 3 & -1 & 5 \\ 2 & 1 & 3 \\ 0 & -5 & 1 \end{bmatrix}$$

Al proceder como en el ejemplo 3.45, se reduce B a su forma escalonada reducida por renglones

$$\begin{bmatrix} 1 & 0 & \frac{8}{5} \\ 0 & 1 & -\frac{1}{5} \\ 0 & 0 & 0 \end{bmatrix}$$

y use los vectores renglón distintos de cero como base para el espacio renglón. Dado que comenzó con vectores columna, debe transponer nuevamente. Por tanto, una base para gen$(\mathbf{u}, \mathbf{v}, \mathbf{w})$ es

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ \frac{8}{5} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -\frac{1}{5} \end{bmatrix} \right\}$$

Comentarios

- De hecho, no es necesario llegar a la forma escalonada *reducida* por renglones, la forma escalonada es suficiente. Si U es una forma escalonada por renglones de A, en-

tonces los vectores renglón distintos de cero de U formarán una base para renglón(A) (vea el ejercicio 33). Este planteamiento tiene la ventaja de (con frecuencia) permitir evitar las fracciones. En el ejemplo 3.46, B puede reducirse a

$$U = \begin{bmatrix} 3 & -1 & 5 \\ 0 & -5 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

que produce la base

$$\left\{ \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}, \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix} \right\}$$

para gen$(\mathbf{u}, \mathbf{v}, \mathbf{w})$.

- Advierta que los métodos usados en el ejemplo 3.44, el ejemplo 3.46 y el comentario anterior por lo general producirán diferentes bases.

Ahora se regresa al problema de encontrar una base para el espacio columna de una matriz A. Un método es simplemente transponer la matriz. Los vectores columna de A se convierten en los vectores renglón de A^T y puede aplicar el método del ejemplo 3.45 para encontrar una base para renglón(A^T). Al transponer estos vectores se produce una base para col(A). (En los ejercicios 21-24 se le pide hacer esto.) Sin embargo, este esquema requiere realizar un nuevo conjunto de operaciones de renglón sobre A^T.

En vez de ello, es preferible un planteamiento que permita usar la forma reducida por renglones de A que ya se calculó. Recuerde que un producto de $A\mathbf{x}$ de una matriz y un vector corresponde a una combinación lineal de las columnas de A con las entradas de $\mathbf{x}$ como coeficientes. Por tanto, una solución no trivial a $A\mathbf{x} = \mathbf{0}$ representa una *relación de dependencia* entre las columnas de A. dado que las operaciones elementales con renglones no afectan al conjunto solución, si A es equivalente por renglones a R, *las columnas de A tienen las mismas relaciones de dependencia que las columnas de R*. Esta importante observación es el fundamento de la técnica que ahora se usa para encontrar una base para col(A).

Ejemplo 3.47

Encuentre una base para el espacio columna de la matriz del ejemplo 3.45,

$$A = \begin{bmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{bmatrix}$$

Solución Sea $\mathbf{a}_i$ un vector columna de A y $\mathbf{r}_i$ un vector columna de la forma escalonada reducida

$$R = \begin{bmatrix} 1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 2 & 0 & 3 \\ 0 & 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Puede ver rápidamente por inspección que $\mathbf{r}_3 = \mathbf{r}_1 + 2\mathbf{r}_2$ y $\mathbf{r}_5 = -\mathbf{r}_1 + 3\mathbf{r}_2 + 4\mathbf{r}_4$. (Compruebe que, como se predijo, los correspondientes vectores columna de A satisfacen las mismas relaciones de dependencia.) Por tanto, $\mathbf{r}_3$ y $\mathbf{r}_5$ no contribuyen a col(R). Los vectores columna restantes, $\mathbf{r}_1$, $\mathbf{r}_2$ y $\mathbf{r}_4$, son linealmente independientes, pues son vectores

unitarios estándar. En consecuencia, los enunciados correspondientes son verdaderos para los vectores columna de A.

Por tanto, entre los vectores columna de A, se eliminan los dependientes ($\mathbf{a}_3$ y $\mathbf{a}_5$), y los restantes serán linealmente independientes y en consecuencia forman una base para col(A). ¿Cuál es la forma más rápida de encontrar esta base? Use las columnas de A que correspondan a las columnas de R *que contengan los 1 pivote.* Una base para col(A) es

$$\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4\} = \left\{ \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -2 \\ 1 \end{bmatrix} \right\}$$

Advertencia ¡Las operaciones elementales con renglones cambian el espacio columna! En el ejemplo, col(A) $\neq$ col(R), ya que todo vector en col(R) tiene su cuarto componente igual a 0, pero esto ciertamente no es verdadero de col(A). De modo que debe regresar a la matriz original A para obtener los vectores columna para una base de col(A). Para ser específico, en el ejemplo 3.47, $\mathbf{r}_1$, $\mathbf{r}_2$ y $\mathbf{r}_4$ *no* forman una base para el espacio columna de A.

Ejemplo 3.48

Encuentre una base para el espacio nulo de la matriz A del ejemplo 3.47.

Solución En realidad no hay nada nuevo aquí, excepto la terminología. Simplemente debe encontrar y describir las soluciones del sistema homogéneo $A\mathbf{x} = 0$. Ya se calculó la forma escalonada reducida por renglones R de A, de modo que todo lo que falta por hacer en la eliminación de Gauss-Jordan es resolver para las variables pivote en términos de las variables libres. La matriz aumentada final es

$$[R \mid \mathbf{0}] = \left[\begin{array}{ccccc|c} 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Si

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$$

entonces los 1 pivote están en las columnas 1, 2 y 4, así que se resuelve para x_1, x_2 y x_4 en términos de las variables libres x_3 y x_5. Se obtiene $x_1 = -x_3 + x_5$, $x_2 = -2x_3 - 3x_5$ y $x_4 = -4x_5$. Al hacer $x_3 = s$ y $x_5 = t$, se obtiene

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -s + t \\ -2s - 3t \\ s \\ -4t \\ t \end{bmatrix} = s\begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix} = s\mathbf{u} + t\mathbf{v}$$

Por tanto, $\mathbf{u}$ y $\mathbf{v}$ generan nul(A), y dado que son linealmente independientes, forman una base para nul(A).

A continuación hay un resumen del procedimiento más efectivo para encontrar bases para el espacio renglón, el espacio columna y el espacio nulo de una matriz A.

1. Encuentre la forma escalonada reducida por renglones R de A.
2. Use los vectores renglón distintos de cero de R (que contengan los 1 pivote) para formar una base para renglón(A).
3. Use los vectores columna de A que correspondan a las columnas de R que contengan los 1 pivote (las columnas pivote) para formar una base para col(A).
4. Resuelva para las variables pivote de $R\mathbf{x} = \mathbf{0}$ en términos de las variables libres, iguale las variables libres a parámetros, sustituya de vuelta en $\mathbf{x}$ y escriba el resultado como una combinación lineal de f vectores (donde f es el número de variables libres). Estos vectores f forman una base para nul(A).

Si no necesita encontrar el espacio nulo, entonces es más rápido simplemente reducir A a la forma escalonada por renglones para encontrar bases para los espacios renglón y columna. Los pasos 2 y 3 anteriores siguen siendo válidos (con la sustitución de la palabra "pivotes" en vez de "1 pivotes").

Dimensión y rank

Ya observó que, aunque un subespacio tendrá diferentes bases, cada una tiene el mismo número de vectores. Este hecho fundamental será de vital importancia en este libro a partir de ahora.

Teorema 3.23 **El Teorema de la Base**

Sea S un subespacio de $\mathbb{R}^n$. Entonces cualesquiera dos bases para S tienen el mismo número de vectores.

Sherlock Holmes apuntó: "Cuando se ha eliminado lo imposible, cualquier cosa que quede, *aunque sea improbable*, debe ser la verdad" (de *El signo de los cuatro*, de sir Arthur Conan Doyle).

Demostración Sean $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ y $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_s\}$ bases de S. Necesita demostrar que $r = s$. Esto se hace al demostrar que no puede ocurrir ninguna de las otras dos posibilidades, $r < s$ o $r > s$.

Suponga que $r < s$. Se demostrará que esto fuerza a $\mathcal{C}$ a ser un conjunto linealmente dependiente de vectores. Para este fin, sea

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_s\mathbf{v}_s = \mathbf{0} \tag{1}$$

Dado que $\mathcal{B}$ es una base para S, puede escribir cada $\mathbf{v}_i$ como una combinación lineal de los elementos $\mathbf{u}_j$:

$$\begin{aligned} \mathbf{v}_1 &= a_{11}\mathbf{u}_1 + a_{12}\mathbf{u}_2 + \cdots + a_{1r}\mathbf{u}_r \\ \mathbf{v}_2 &= a_{21}\mathbf{u}_1 + a_{22}\mathbf{u}_2 + \cdots + a_{2r}\mathbf{u}_r \\ &\vdots \\ \mathbf{v}_s &= a_{s1}\mathbf{u}_1 + a_{s2}\mathbf{u}_2 + \cdots + a_{sr}\mathbf{u}_r \end{aligned} \tag{2}$$

Al sustituir las ecuaciones (2) en la ecuación (1) se obtiene

$$c_1(a_{11}\mathbf{u}_1 + \cdots + a_{1r}\mathbf{u}_r) + c_2(a_{21}\mathbf{u}_1 + \cdots + a_{2r}\mathbf{u}_r) + \cdots + c_s(a_{s1}\mathbf{u}_1 + \cdots + a_{sr}\mathbf{u}_r) = \mathbf{0}$$

Al reagrupar se tiene

$$(c_1a_{11} + c_2a_{21} + \cdots + c_sa_{s1})\mathbf{u}_1 + (c_1a_{12} + c_2a_{22} + \cdots + c_sa_{s2})\mathbf{u}_2 + \cdots + (c_1a_{1r} + c_2a_{2r} + \cdots + c_sa_{sr})\mathbf{u}_r = \mathbf{0}$$

Ahora, dado que $\mathcal{B}$ es una base, los términos $\mathbf{u}_j$ son linealmente independientes. De modo que cada una de las expresiones entre paréntesis debe ser cero:

$$\begin{aligned} c_1a_{11} + c_2a_{21} + \cdots + c_sa_{s1} &= 0 \\ c_1a_{12} + c_2a_{22} + \cdots + c_sa_{s2} &= 0 \\ &\vdots \\ c_1a_{1r} + c_2a_{2r} + \cdots + c_sa_{sr} &= 0 \end{aligned}$$

Este es un sistema homogéneo de r ecuaciones lineales con las s variables $c_1, c_2, \ldots, c_s$. (El hecho de que las variables aparezcan a la izquierda de los coeficientes no hace diferencia.) Dado que $r < s$, se sabe del Teorema 2.3 que existe un número infinito de soluciones. En particular, existe una solución no trivial, lo que da una relación de dependencia no trivial en la ecuación (1). Por tanto, $\mathcal{C}$ es un conjunto linealmente dependiente de vectores. Pero este hallazgo contradice el hecho de que $\mathcal{C}$ fue dado como base y en consecuencia es linealmente independiente. Se concluye que $r < s$ no es posible. De igual modo (al intercambiar los roles de $\mathcal{B}$ y $\mathcal{C}$), se descubre que $r > s$ conduce a una contradicción. Por tanto, debe tener $r = s$, como se deseaba.

Dado que todas las bases para un subespacio dado deben tener el mismo número de vectores, puede darse un nombre a este número.

Definición Si S es un subespacio de $\mathbb{R}^n$, entonces el número de vectores en una base para S se llama la ***dimensión*** de S, y se denota dim S.



Comentario El vector cero $\mathbf{0}$ por sí mismo siempre es un subespacio de $\mathbb{R}^n$. (¿Por qué?) Sin embargo, cualquier conjunto que contenga el vector cero (y, en particular, $\{\mathbf{0}\}$) es linealmente dependiente, de modo que $\{\mathbf{0}\}$ no puede tener una base. Se define $\{\mathbf{0}\}$ como 0.

Ejemplo 3.49 Dado que la base estándar para $\mathbb{R}^n$ tiene n vectores, $\mathbb{R}^n = n$. (Note que este resultado concuerda con la comprensión intuitiva de dimensión para $n \leq 3$.)

Ejemplo 3.50 En los ejemplos 3.45 a 3.48, se descubrió que renglón(A) tiene una base con tres vectores, col(A) tiene una base con tres vectores y nul(A) tiene una base con dos vectores. En consecuencia, (renglón(A)) = 3, dim (col(A)) = 3 y dim (nul(A)) = 2.

Un solo ejemplo no es suficiente para especular, pero el hecho de que los espacios renglón y columna del ejemplo 3.50 tengan la misma dimensión no es accidente. Como tampoco lo es el hecho de que la suma de dim(col(A)) y dim(nul(A)) es 5, el número de columnas de A. Ahora se probará que estas relaciones son verdaderas en general.

Teorema 3.24 Los espacios renglón y columna de una matriz A tienen la misma dimensión.

Demostración Sea R la forma escalonada reducida por renglones de A. Por el Teorema 3.20, renglón(A) = renglón(R), de modo que

$$\begin{aligned}\dim(\text{renglón}(A)) &= \dim(\text{renglón}(R))\\ &= \text{número de renglones de } R \text{ distintos de cero}\\ &= \text{número de 1 pivote de } R\end{aligned}$$

Llame a este número r.

Ahora col$(A) \neq$ col(R), pero las columnas de A y R tienen las mismas relaciones de dependencia. Por tanto, dim (col(A)) = dim (col(R)). Dado que existen r 1 pivotes, R tiene r columnas que son vectores unitarios estándar, $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$. (Éstos serán vectores en $\mathbb{R}^m$ si A y R son matrices de $m \times n$.) Esos r vectores son linealmente independientes y las columnas restantes de R son combinaciones lineales de ellos. Por tanto, dim (col(R)) $= r$. Se tiene que dim (renglón(A)) $= r =$ dim (col(A)), como se quería demostrar.

El rango de una matriz la definió por primera vez, en 1878, Georg Frobenius (1849–1917), aunque él la definió usando determinantes y no como se hace aquí. (Vea el capítulo 4.) Frobenius fue un matemático alemán que recibió su doctorado en la Universidad de Berlín, donde más tarde dio clases. Mejor conocido por sus aportaciones a la teoría de grupos, Frobenius usó matrices en su trabajo con representaciones de grupo.

Definición El ***rank*** de una matriz A es la dimensión de sus espacios renglón y columna y se denota mediante rango(A).

Por tanto, para el ejemplo 3.50, puede escribir rango$(A) = 3$.

Comentarios

- La definición anterior concuerda con la definición más informal de rank que se dio en el capítulo 2. La ventaja de la nueva definición es que es mucho más flexible.
- El rank de una matriz simultáneamente proporciona información acerca de la dependencia lineal entre los vectores renglón de la matriz *y* entre sus vectores columna. En particular, dice el número de renglones y columnas que son linealmente independientes (¡y este número es el mismo en cada caso!).

Dado que los vectores renglón de A son los vectores columna de A^T, el Teorema 3.24 tiene el siguiente corolario inmediato.

Teorema 3.25 Para cualquier matriz A,

$$\text{rank}(A^T) = \text{rank}(A)$$

Demostración Se tiene

$$\begin{aligned}\text{rank}(A^T) &= \dim(\text{col}(A^T))\\ &= \dim(\text{renglón}(A))\\ &= \text{rank } A\end{aligned}$$

Definición La ***nulidad*** de una matriz A es la dimensión de su espacio nulo y se denota por nulidad(A).

En otras palabras, nulidad(A) es la dimensión del espacio solución de $A\mathbf{x} = \mathbf{0}$, que es igual que el número de variables libres en la solución. Ahora puede regresar al teorema del rank (Teorema 2.2) y parafrasearlo en términos de las nuevas definiciones.

Teorema 3.26 **El Teorema del rank**

Si A es una matriz de $m \times n$, entonces

$$\text{rank}(A) + \text{nulidad}(A) = n$$

Demostración Sea R la forma escalonada reducida por renglones de A, y suponga que $\text{rank}(A) = r$. Entonces R tiene r 1 pivote, de modo que existen r variables pivote y $n - r$ variables libres en la solución de $A\mathbf{x} = \mathbf{0}$. Puesto que $\dim(\text{nul}(A)) = n - r$, se tiene

$$\begin{aligned}\text{rank}(A) + \text{nulidad}(A) &= r + (n - r)\\ &= n\end{aligned}$$

Con frecuencia, cuando se necesita conocer la nulidad de una matriz, no es necesario conocer la solución real de $A\mathbf{x} = \mathbf{0}$. El teorema del rank es extremadamente útil en tales situaciones, como ilustra el siguiente ejemplo.

Ejemplo 3.51

Encuentre la nulidad de cada una de las siguientes matrices:

$$M = \begin{bmatrix} 2 & 3 \\ 1 & 5 \\ 4 & 7 \\ 3 & 6 \end{bmatrix} \quad \text{y}$$

$$N = \begin{bmatrix} 2 & 1 & -2 & -1 \\ 4 & 4 & -3 & 1 \\ 2 & 7 & 1 & 8 \end{bmatrix}$$

Solución Dado que las dos columnas de M claramente son linealmente independientes, $\text{rank}(M) = 2$. Por tanto, por el teorema del rank, $\text{nulidad}(M) = 2 - \text{rank}(M) = 2 - 2 = 0$.

No hay dependencia obvia entre los renglones o columnas de N, así que se aplican operaciones con renglones para reducirla a

$$\begin{bmatrix} 2 & 1 & -2 & -1 \\ 0 & 2 & 1 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La matriz se redujo lo suficiente (aquí no es necesaria la forma escalonada *reducida* por renglones, pues no se busca una base para el espacio nulo). Se ve que sólo existen dos renglones distintos de cero, de modo que $\text{rank}(N) = 2$. En consecuencia, nulidad $(N) = 4 - \text{rank}(N) = 4 - 2 = 2$.

Los resultados de estas acciones permiten extender el teorema fundamental de las matrices invertibles (Teorema 3.12).

Teorema 3.27 **El teorema fundamental de las matrices invertibles: versión 2**

Sea A una matriz de $n \times n$. Los siguientes enunciados son equivalentes:

a. A es invertible.
b. $A\mathbf{x} = \mathbf{b}$ tiene una solución única para todo $\mathbf{b}$ en $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial.
d. La forma escalonada reducida por renglones de A es I_n.
e. A es un producto de matrices elementales.
f. $\text{rank}(A) = n$
g. $\text{nulidad}(A) = 0$
h. Los vectores columna de A son linealmente independientes.
i. Los vectores columna de A generan $\mathbb{R}^n$.
j. Los vectores columna de A forman una base para $\mathbb{R}^n$.
k. Los vectores renglón de A son linealmente independientes.
l. Los vectores renglón de A generan $\mathbb{R}^n$.
m. Los vectores renglón de A forman una base para $\mathbb{R}^n$.

La nulidad de una matriz la definió en 1884 James Joseph Sylvester (1814–1887), quien estaba interesado en las *invariantes*: propiedades de las matrices que no cambian en ciertos tipos de transformaciones. Originario de Inglaterra, Sylvester se convirtió en el segundo presidente de la Sociedad Matemática de Londres. En 1878, mientras impartía clases en la Johns Hopkins University en Baltimore, fundó el *American Journal of Mathematics*, la primera revista matemática en Estados Unidos.

Demostración Ya se estableció la equivalencia desde (a) hasta (e). Falta por demostrar que los enunciados (f) al (m) son equivalentes a los primeros cinco enunciados.

(f) $\Leftrightarrow$ (g) Dado que $\text{rank}(A) + \text{nulidad}(A) = n$ cuando A es una matriz de $n \times n$, se tiene del teorema del rank que $\text{rank}(A) = n$ si y sólo si $\text{nulidad}(A) = 0$.

(f) $\Rightarrow$ (d) $\Rightarrow$ (c) $\Rightarrow$ (h) Si $\text{rank}(A) = n$, entonces la forma escalonada reducida por renglón de A tiene n 1 pivotes, y por tanto es I_n. De (d) $\Rightarrow$ (c) se sabe que $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial, lo que implica que los vectores columna de A son linealmente independientes, pues $A\mathbf{x}$ es justo una combinación lineal de los vectores columna de A.

(h) $\Rightarrow$ (i) Si los vectores columna de A son linealmente independientes, entonces $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial. Por tanto, por (c) $\Rightarrow$ (b), $A\mathbf{x} = \mathbf{b}$, tiene una solución única para todo $\mathbf{b}$ en $\mathbb{R}^n$. Esto significa que todo vector $\mathbf{b}$ en $\mathbb{R}^n$ puede escribirse como una combinación lineal de los vectores columna de A, lo que establece (i)

(i) $\Rightarrow$ (j) Si los vectores columna de A generan $\mathbb{R}^n$, entonces $\text{col}(A) = \mathbb{R}^n$ por definición, de modo que $\text{rank}(A) = \dim(\text{col}(A)) = n$. Esto es (f), y ya se estableció que (f) $\Rightarrow$ (h). Se concluye que los vectores columna de A son linealmente independientes y por tanto forman una base para $\mathbb{R}^n$, pues, por suposición, también generan $\mathbb{R}^n$.

(j) $\Rightarrow$ (f) Si los vectores columna de A forman una base para $\mathbb{R}^n$, entonces, en particular, son linealmente independientes. Se tiene que la forma escalonada reducida por renglones de A contiene n 1 pivotes, y por tanto $\text{rank}(A) = n$.

El análisis anterior demuestra que (f) $\Rightarrow$ (d) $\Rightarrow$ (c) $\Rightarrow$ (h) $\Rightarrow$ (i) $\Rightarrow$ (j) $\Rightarrow$ (f) $\Leftrightarrow$ (g). Ahora recuerde que, por el Teorema 3.25, $\text{rank}(A^T) = \text{rank}(A)$, de modo que lo recién demostrado da los resultados correspondientes acerca de los vectores columna de A^T. Estos son entonces resultados acerca de los vectores *renglón* de A, lo que incluye en (k), (l) y (m) a la red de equivalencias y completa la demostración.

Los teoremas como el teorema fundamental no son simplemente de interés teórico. También son tremendos dispositivos que ahorran trabajo. El teorema fundamental ya permitió recortar a la mitad el trabajo necesario para comprobar que dos matrices cuadradas son inversas. También simplifica la tarea de demostrar que ciertos conjuntos de vectores son bases para $\mathbb{R}^n$. De hecho, cuando se tiene un conjunto de n vectores en $\mathbb{R}^n$, dicho conjunto será una base para $\mathbb{R}^n$ si es verdadera *alguna* de las propiedades necesarias de independencia lineal o de conjunto generador. El siguiente ejemplo muestra cuán fáciles pueden ser los cálculos.

Ejemplo 3.52

Demuestre que los vectores

$$\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}-1\\0\\1\end{bmatrix} \text{ y } \begin{bmatrix}4\\9\\7\end{bmatrix}$$

forman una base para $\mathbb{R}^3$.

Solución De acuerdo con el Teorema Fundamental, los vectores formarán una base para $\mathbb{R}^3$ si y sólo si una matriz con estos vectores como sus columnas (o renglones) tiene rank 3. Se realizan sólo suficientes operaciones con renglones para determinar esto:

$$A = \begin{bmatrix}1 & -1 & 4\\2 & 0 & 9\\3 & 1 & 7\end{bmatrix} \longrightarrow \begin{bmatrix}1 & -1 & 4\\0 & 2 & 1\\0 & 0 & -7\end{bmatrix}$$

Se ve que A tiene rank 3, de modo que los vectores dados son una base para $\mathbb{R}^3$, por la equivalencia de (f) y (j).

El siguiente teorema es una aplicación tanto del teorema del rank como del teorema fundamental. Este resultado se requerirá en los capítulos 5 y 7.

Teorema 3.28 Sea A una matriz de $m \times n$. Entonces:

a. $\text{rank}(A^TA) = \text{rank}(A)$
b. La matriz A^TA de $n \times n$ es invertible si y sólo si $\text{rank}(A) = n$.

Demostración
(a) Dado que A^TA es de $n \times n$, tiene el mismo número de columnas que A. El teorema del rank dice entonces que

$$\text{rank}(A) + \text{nulidad}(A) = n = \text{rank}(A^TA) + \text{nulidad}(A^TA)$$

Por tanto, para demostrar que $\text{rank}(A) = \text{rank}(A^TA)$, es suficiente demostrar que $\text{nulidad}(A) = \text{nulidad}(A^TA)$. Se hará así al establecer que los espacios nulos de A y A^TA son iguales.

Para este fin, sea $\mathbf{x}$ en $\text{nul}(A)$, de modo que $A\mathbf{x} = \mathbf{0}$. Entonces $A^TA\mathbf{x} = A^T\mathbf{0} = \mathbf{0}$, y por tanto $\mathbf{x}$ está en $\text{nul}(A^TA)$. A la inversa, sea $\mathbf{x}$ en $\text{nul}(A^TA)$. Entonces $A^TA\mathbf{x} = \mathbf{0}$, de modo que $\mathbf{x}^TA^TA\mathbf{x} = \mathbf{x}^T\mathbf{0} = 0$. Pero entonces

$$(A\mathbf{x}) \cdot (A\mathbf{x}) = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^TA^TA\mathbf{x} = 0$$

y en consecuencia $A\mathbf{x} = 0$, por el Teorema 1.2(d). Por tanto, $\mathbf{x}$ está en $\text{nul}(A)$, de modo que $\text{nul}(A) = \text{nul}(A^TA)$, como se requería.

(b) Por el teorema fundamental, la matriz A^TA de $n \times n$ es invertible si y sólo si $\text{rank}(A^TA) = n$. Pero, por (a), esto se cumple si y sólo si $\text{rank}(A) = n$.

Coordenadas

Ahora regrese a una de las preguntas planteadas al comienzo de esta sección: ¿cómo se verían los vectores en $\mathbb{R}^3$ que existen en un plano que pasa por el origen? ¿Son bidimensionales o tridimensionales? Las nociones de base y dimensión ayudarán a clarificar las cosas.

Un plano que pasa por el origen es un subespacio bidimensional de $\mathbb{R}^3$, con cualquier conjunto de dos vectores dirección que funcionan como base. Los vectores base ubican ejes coordenados en el plano-subespacio y a su vez permiten ver el plano como una "copia" de $\mathbb{R}^2$. Antes de ilustrar este enfoque, se demostrará un teorema que garantiza que las "coordenadas" surgidas de esta forma son únicas.

Teorema 3.29 Sea S un subespacio de $\mathbb{R}^n$ y sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base para S. Para todo vector $\mathbf{v}$ en S, existe exactamente una forma de escribir $\mathbf{v}$ como una combinación lineal de los vectores base en $\mathcal{B}$:

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

Demostración Dado que $\mathcal{B}$ es una base, genera a S, de modo que $\mathbf{v}$ puede escribirse en *al menos una* forma como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Sea una de dichas combinaciones lineales

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

Su tarea es demostrar que ésta es la *única* forma de escribir $\mathbf{v}$ como combinación lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Para este fin, suponga que también se tiene

$$\mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k$$

Entonces $\quad c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_k\mathbf{v}_k$

Al reordenar (con las propiedades del álgebra vectorial), se obtiene

$$(c_1 - d_1)\mathbf{v}_1 + (c_2 - d_2)\mathbf{v}_2 + \cdots + (c_k - d_k)\mathbf{v}_k = \mathbf{0}$$

Dado que $\mathcal{B}$ es una base, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ son linealmente independientes. Por tanto,

$$(c_1 - d_1) = (c_2 - d_2) = \cdots = (c_k - d_k) = 0$$

En otras palabras, $c_1 = d_1, c_2 = d_2, \ldots, c_k = d_k$, y las dos combinaciones lineales en realidad son iguales. Por ende, existe exactamente una forma de escribir $\mathbf{v}$ como una combinación lineal de los vectores base en $\mathcal{B}$.

Definición Sea S un subespacio de $\mathbb{R}^n$ y sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base para S. Sea $\mathbf{v}$ un vector en S, y escriba $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$. Entonces $c_1, c_2, \ldots, c_k$ se llaman las ***coordenadas de* v *con respecto a* $\mathcal{B}$**, y el vector columna

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix}$$

se llama ***vector coordenada de* v *con respecto a* $\mathcal{B}$.**

Ejemplo 3.53 Sea $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ la base estándar de $\mathbb{R}^3$. Encuentre el vector coordenada de

$$\mathbf{v} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}$$

con respecto a $\mathcal{E}$.

Solución Dado que $\mathbf{v} = 2\mathbf{e}_1 + 7\mathbf{e}_2 + 4\mathbf{e}_3$,

$$[\mathbf{v}]_{\mathcal{E}} = \begin{bmatrix} 2 \\ 7 \\ 4 \end{bmatrix}$$

Debe ser claro que el vector coordenada de todo vector (columna) en $\mathbb{R}^n$ con respecto a la base estándar es justo el vector mismo.

Ejemplo 3.54

En el ejemplo 3.44 se vio que $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix}$, $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, y $\mathbf{w} = \begin{bmatrix} 0 \\ -5 \\ 1 \end{bmatrix}$ son tres vectores en el mismo subespacio (plano que pasa por el origen) S de $\mathbb{R}^3$ y que $\mathcal{B} = \{\mathbf{u}, \mathbf{v}\}$ es una base para S. Dado que $\mathbf{w} = 2\mathbf{u} - 3\mathbf{v}$, se tiene

$$[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$$

Vea la figura 3.3.

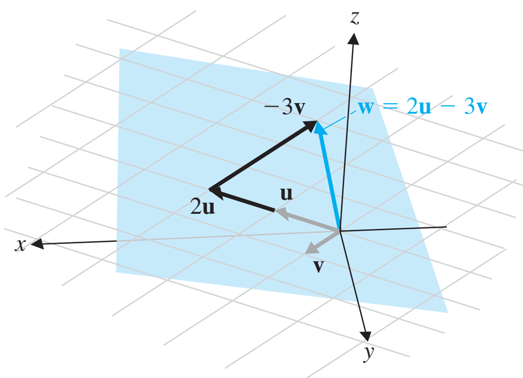


Figura 3.3
Las coordenadas de un vector con respecto a una base

Ejercicios 3.5

En los ejercicios 1-4, sea S la colección de vectores $\begin{bmatrix} x \\ y \end{bmatrix}$ *en* $\mathbb{R}^2$ *que satisfacen la propiedad dada. En cada caso, demuestre que S forma un subespacio de* $\mathbb{R}^2$ *u ofrezca un contraejemplo para demostrar que no lo hace.*

1. $x = 0$ **2.** $x \geq 0, y \geq 0$

3. $y = 2x$ **4.** $xy \geq 0$

En los ejercicios 5-8, sea S la colección de vectores $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ *en* $\mathbb{R}^3$ *que satisfacen la propiedad dada. En cada caso, demuestre que S forma un subespacio de* $\mathbb{R}^3$ *u ofrezca un contraejemplo para demostrar que no lo hace.*

5. $x = y = z$ **6.** $z = 2x, y = 0$

7. $x - y + z = 1$ **8.** $|x - y| = |y - z|$

9. Demuestre que toda recta que pasa por el origen en $\mathbb{R}^3$ es un subespacio de $\mathbb{R}^3$.

10. Suponga que S consiste de todos los puntos en $\mathbb{R}^2$ que están sobre el eje x o el eje y (o ambos). (S se llama la *unión* de los dos ejes.) ¿S es un subespacio de $\mathbb{R}^2$? ¿Por qué sí o por qué no?

En los ejercicios 11 y 12, determine si **b** *está en col(A) y si* **w** *está en renglón(A), como en el ejemplo 3.41.*

11. $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $\mathbf{w} = [-1 \quad 1 \quad 1]$

12. $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 3 & 0 \\ 3 & -1 & -5 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{w} = [1 \quad -3 \quad -3]$

13. En el ejercicio 11, determine si $\mathbf{w}$ está en renglón(A), usando el método descrito en el comentario que sigue al ejemplo 3.41.

14. En el ejercicio 12, determine si $\mathbf{w}$ está en renglón(A), usando el método descrito en el comentario que sigue al ejemplo 3.41.

15. Si A es la matriz del ejercicio 11, ¿$\mathbf{v} = \begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}$ está en nul(A)?

16. Si A es la matriz del ejercicio 12, ¿$\mathbf{v} = \begin{bmatrix} 7 \\ -1 \\ 2 \end{bmatrix}$ está en nul(A)?

En los ejercicios 17-20, proporcione bases para renglón(A), col(A) y nul(A).

17. $A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{bmatrix}$ **18.** $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 5 & 1 \\ 1 & -1 & -2 \end{bmatrix}$

19. $A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & -1 \end{bmatrix}$

20. $A = \begin{bmatrix} 2 & -4 & 0 & 2 & 1 \\ -1 & 2 & 1 & 2 & 3 \\ 1 & -2 & 1 & 4 & 4 \end{bmatrix}$

En los ejercicios 21-24, encuentre bases para renglón(A) y col(A) en los ejercicios dados usando A^T.

21. Ejercicio 17 **22.** Ejercicio 18

23. Ejercicio 19 **24.** Ejercicio 20

25. Explique cuidadosamente por qué sus respuestas a los ejercicios 17 y 21 son ambas correctas, aun cuando parezca haber diferencias.

26. Explique cuidadosamente por qué sus respuestas a los ejercicios 18 y 22 son ambas correctas, aun cuando parezca haber diferencias.

En los ejercicios 27-30, encuentre una base para el generador de los vectores dados.

27. $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ **28.** $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

29. $[2 \quad -3 \quad 1], [1 \quad -1 \quad 0], [4 \quad -4 \quad 1]$

30. $[3 \quad 1 \quad -1 \quad 0], [0 \quad -1 \quad 2 \quad -1], [4 \quad 3 \quad 8 \quad 3]$

Para los ejercicios 31 y 32, encuentre bases para los generadores de los vectores en los ejercicios dados de entre los vectores mismos.

31. Ejercicio 29 **32.** Ejercicio 30

33. Demuestre que si R es una matriz en forma escalonada, entonces una base para renglón(R) consiste de los renglones distintos de cero de R.

34. Demuestre que si las columnas de A son linealmente independientes, entonces deben formar una base para col(A).

Para los ejercicios 35-38, proporcione el rank y la nulidad de las matrices en los ejercicios dados.

35. Ejercicio 17

36. Ejercicio 18

37. Ejercicio 19

38. Ejercicio 20

39. Si A es una matriz de 3×5, explique por qué las columnas de A deben ser linealmente dependientes.

40. Si A es una matriz de 4×2, explique por qué los renglones de A deben ser linealmente dependientes.

41. Si A es una matriz de 3×5, ¿cuáles son los posibles valores de nulidad(A)?

42. Si A es una matriz de 4×2, ¿cuáles son los posibles valores de nulidad(A)?

En los ejercicios 43 y 44, encuentre todos los posibles valores de rank(A) cuando a varía.

43. $A = \begin{bmatrix} 1 & 2 & a \\ -2 & 4a & 2 \\ a & -2 & 1 \end{bmatrix}$ **44.** $A = \begin{bmatrix} a & 2 & -1 \\ 3 & 3 & -2 \\ -2 & -1 & a \end{bmatrix}$

Responda los ejercicios 45-48 al considerar la matriz con los vectores dados como sus columnas.

45. ¿$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$ forma una base para $\mathbb{R}^3$?

46. ¿$\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -5 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 4 \end{bmatrix}$ forma una base para $\mathbb{R}^3$?

47. ¿$\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ forma una base para $\mathbb{R}^4$?

48. ¿ $\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\-1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\-1\end{bmatrix}$ forma una base para $\mathbb{R}^4$?

49. ¿ $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}$ forma una base para $\mathbb{Z}_2^3$?

50. ¿ $\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}$ forma una base para $\mathbb{Z}_3^3$?

En los ejercicios 51 y 52, demuestre que **w** *está en gen($\mathcal{B}$) y encuentre el vector coordenada* $[\mathbf{w}]_{\mathcal{B}}$.

51. $\mathcal{B} = \left\{\begin{bmatrix}1\\2\\0\end{bmatrix}, \begin{bmatrix}1\\0\\-1\end{bmatrix}\right\}, \mathbf{w} = \begin{bmatrix}1\\6\\2\end{bmatrix}$

52. $\mathcal{B} = \left\{\begin{bmatrix}3\\1\\4\end{bmatrix}, \begin{bmatrix}5\\1\\6\end{bmatrix}\right\}, \mathbf{w} = \begin{bmatrix}1\\3\\4\end{bmatrix}$

En los ejercicios 53-56, calcule el rank y la nulidad de las matrices dadas sobre el $\mathbb{Z}_p$ indicado.

53. $\begin{bmatrix}1&1&0\\0&1&1\\1&0&1\end{bmatrix}$ sobre $\mathbb{Z}_2$ **54.** $\begin{bmatrix}1&1&2\\2&1&2\\2&0&0\end{bmatrix}$ sobre $\mathbb{Z}_3$

55. $\begin{bmatrix}1&3&1&4\\2&3&0&1\\1&0&4&0\end{bmatrix}$ sobre $\mathbb{Z}_5$

56. $\begin{bmatrix}2&4&0&0&1\\6&3&5&1&0\\1&0&2&2&5\\1&1&1&1&1\end{bmatrix}$ sobre $\mathbb{Z}_7$

57. Si A es de $m \times n$, demuestre que todo vector en nul(A) es ortogonal a todo vector en renglón(A).

58. Si A y B son matrices de $n \times n$ de rank n, demuestre que AB tiene rank n.

59. **(a)** Demuestre que rank(AB) $\leq$ rank(B). [*Sugerencia:* revise el ejercicio 29 de la sección 3.1.]
(b) Ofrezca un ejemplo en el que rank(AB) $<$ rank(B).

60. **(a)** Demuestre que rank(AB) $\leq$ rank(A). [*Sugerencia:* Revise el ejercicio 30 de la sección 3.1 o use transpuestas y el ejercicio 59(a).]
(b) Ofrezca un ejemplo en el que rank(AB) $=$ rank(A).

61. **(a)** Demuestre que si U es invertible, entonces rank(UA) $=$ rank(A). [*Sugerencia:* $A = U^{-1}(UA)$.]
(b) Demuestre que si V es invertible, entonces rank(AV) $=$ rank(A).

62. Demuestre que una matriz A de $m \times n$ tiene rank 1 si y sólo si A puede escribirse como el producto exterior $\mathbf{u}\mathbf{v}^T$ de un vector $\mathbf{u}$ en $\mathbb{R}^m$ y $\mathbf{v}$ en $\mathbb{R}^n$.

63. Si una matriz A de $m \times n$ tiene rank r, demuestre que A puede escribirse como la suma de r matrices, cada una de las cuales tiene rank 1. [*Sugerencia:* encuentre una forma de usar el ejercicio 62.]

64. Demuestre que, para matrices A y B de $m \times n$, rank($A + B$) $\leq$ rank(A) $+$ rank(B).

65. Sea A una matriz $n \times n$ tal que $A^2 = O$. Demuestre que rank(A) $\leq n/2$. [*Sugerencia:* demuestre que col(A) $\subseteq$ nul(A) y use el teorema del rank.]

66. Sea A una matriz antisimétrica $n \times n$. (Vea la página 168.)
(a) Demuestre que $\mathbf{x}^T A\mathbf{x} = 0$ para todo $\mathbf{x}$ en $\mathbb{R}^n$.
(b) Demuestre que $I + A$ es invertible. [*Sugerencia:* demuestre que nul($I + A$) $= \{\mathbf{0}\}$.]

3.6 Introducción a las transformaciones lineales

En esta sección comienza la exploración de uno de los temas de la introducción a este capítulo. Ahí se vio que la matrices pueden usarse para transformar vectores y actuar como un tipo de "función" de la forma $\mathbf{w} = T(\mathbf{v})$, donde la variable independiente $\mathbf{v}$ y la variable dependiente $\mathbf{w}$ son vectores. Ahora esta noción se hará más precisa y se observarán varios ejemplos de tales transformaciones matriciales, lo que conduce al concepto de *transformación lineal*, una poderosa idea que se encontrará repetidamente a partir de ahora.

Comience por recordar algunos de los conceptos básicos asociados con las funciones. Se familiarizará con más de estas ideas en otros cursos en los que encontrará funciones de la forma $f: \mathbb{R} \to \mathbb{R}$ [como $f(x) = x^2$] que transforman números reales en números reales. Lo que es novedoso aquí es que están involucrados vectores y que sólo se está interesado en funciones que son "compatibles" con las operaciones vectoriales de suma y multiplicación escalar.

Considere un ejemplo. Sean

$$A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Entonces

$$A\mathbf{v} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$$

Esto demuestra que A transforma $\mathbf{v}$ en $\mathbf{w} = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$.

Esta transformación puede describirse de manera más general. La ecuación matricial

$$\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

proporciona una fórmula que muestra cómo A transforma un vector arbitrario $\begin{bmatrix} x \\ y \end{bmatrix}$ en $\mathbb{R}^2$ en el vector $\begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$ en $\mathbb{R}^3$. Esta transformación se denota T_A y se escribe

$$T_A\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

(Aunque técnicamente desordenada, omitir los paréntesis en definiciones como esta es una convención común que ahorra cierta escritura. La descripción de T_A se convierte en

$$T_A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

con esta convención.)

Con este ejemplo en mente, ahora considere alguna terminología. Una ***transformación*** (o ***mapeo*** o ***función***) T de $\mathbb{R}^n$ en $\mathbb{R}^m$ es una regla que asigna a cada vector $\mathbf{v}$ en $\mathbb{R}^n$ un vector único $T(\mathbf{v})$ en $\mathbb{R}^m$. El ***dominio*** de T es $\mathbb{R}^n$, y el ***codominio*** de T es $\mathbb{R}^m$. Esto se indica al escribir $T: \mathbb{R}^n \to \mathbb{R}^m$. Para un vector $\mathbf{v}$ en el dominio de T, el vector $T(\mathbf{v})$ en el codominio se conoce como la ***imagen*** de $\mathbf{v}$ bajo (la acción de) T. Al conjunto de todas las posibles imágenes $T(\mathbf{v})$ (conforme $\mathbf{v}$ varía a lo largo del dominio de T) se le llama ***rango*** de T.

En el ejemplo, el dominio de T_A es $\mathbb{R}^2$ y su codominio es $\mathbb{R}^3$, así que se escribe $T_A: \mathbb{R}^2 \to \mathbb{R}^3$. La imagen de $\mathbf{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ es $\mathbf{w} = T_A(\mathbf{v}) = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$. Cuál es el rango de T_A?

Consiste de todos los vectores en el codominio $\mathbb{R}^3$ que son de la forma

$$T_A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$$

que describe al conjunto de todas las combinaciones lineales de los vectores columna $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}$ de A. En otras palabras, ¡el rango de T es el espacio columna de A! (Más adelante se hablará más acerca de esto; por el momento, simplemente se indicará como una comentario interesante.) Geométricamente, esto demuestra que el rango de T_A es el plano que pasa por el origen en $\mathbb{R}^3$ con vectores directores dados por los vectores columna de A. Note que el rango de T_A es estrictamente menor que el codominio de T_A.

Transformaciones lineales

El ejemplo T_A anterior es un caso especial de un tipo más general de transformación llamada *transformación lineal*. En el capítulo 6 se considerará la definición general, pero la esencia de ella es que se trata de transformaciones que "conservan" las operaciones vectoriales de suma y multiplicación por un escalar.

Definición Una transformación $T: \mathbb{R}^n \to \mathbb{R}^m$ se llama ***transformación lineal*** si

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ para todo $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$ y
2. $T(c\mathbf{v}) = cT(\mathbf{v})$ para todo $\mathbf{v}$ en $\mathbb{R}^n$ y todo escalar c.

Ejemplo 3.55 Considere una vez más la transformación $T: \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix}$$

Compruebe que T es una transformación lineal. Para verificar (1), sea

$$\mathbf{u} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$$

Entonces

$$T(\mathbf{u} + \mathbf{v}) = T\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) = T\left(\begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 \\ 2(x_1 + x_2) - (y_1 + y_2) \\ 3(x_1 + x_2) + 4(y_1 + y_2) \end{bmatrix}$$

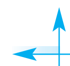
$$= \begin{bmatrix} x_1 + x_2 \\ 2x_1 + 2x_2 - y_1 - y_2 \\ 3x_1 + 3x_2 + 4y_1 + 4y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ (2x_1 - y_1) + (2x_2 - y_2) \\ (3x_1 + 4y_1) + (3x_2 + 4y_2) \end{bmatrix}$$

$$= \begin{bmatrix} x_1 \\ 2x_1 - y_1 \\ 3x_1 + 4y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ 2x_2 - y_2 \\ 3x_2 + 4y_2 \end{bmatrix} = T\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + T\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = T(\mathbf{u}) + T(\mathbf{v})$$

Para demostrar (2), sea $\mathbf{v} = \begin{bmatrix} x \\ y \end{bmatrix}$ y sea c un escalar. Entonces

$$\begin{aligned} T(c\mathbf{v}) &= T\left(c\begin{bmatrix} x \\ y \end{bmatrix}\right) = T\left(\begin{bmatrix} cx \\ cy \end{bmatrix}\right) \\ &= \begin{bmatrix} cx \\ 2(cx) - (cy) \\ 3(cx) + 4(cy) \end{bmatrix} = \begin{bmatrix} cx \\ c(2x - y) \\ c(3x + 4y) \end{bmatrix} \\ &= c\begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = cT\begin{bmatrix} x \\ y \end{bmatrix} = cT(\mathbf{v}) \end{aligned}$$

Por tanto, T es una transformación lineal.



Comentario La definición de una transformación lineal puede simplificarse al combinar (1) y (2) como se muestra a continuación.

$T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal si

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) \quad \text{para todo } \mathbf{v}_1, \mathbf{v}_2 \text{ en } \mathbb{R}^n \text{ y escalares } c_1, c_2$$

En el ejercicio 53 se le pedirá demostrar que el enunciado anterior es equivalente a la definición original. En la práctica, esta formulación equivalente puede ahorrar alguna escritura, ¡inténtelo!

Aunque la transformación lineal T en el ejemplo 3.55 originalmente surgió como una transformación *matricial* T_A, es asunto sencillo recuperar la matriz A a partir de la definición de T dada en el ejemplo. Observe que

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 2x - y \\ 3x + 4y \end{bmatrix} = x\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

de modo que $T = T_A$, donde $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}$. (Note que, cuando las variables x y y se alinean, la matriz A es justo su matriz de coeficientes.)

Es importante reconocer que una transformación es una transformación matricial, pues, como muestra el siguiente teorema, todas las transformaciones matriciales son transformaciones lineales.

Teorema 3.30 Sea A una matriz de $m \times n$. Entonces la transformación matricial $T_A : \mathbb{R}^n \to \mathbb{R}^m$ definida por

$$T_A(\mathbf{x}) = A\mathbf{x} \qquad (\text{para } \mathbf{x} \text{ en } \mathbb{R}^n)$$

es una transformación lineal.

Demostración Sean **u** y **v** vectores en $\mathbb{R}^n$ y sea c un escalar. Entonces

$$T_A(\mathbf{u} + \mathbf{v}) = A(\mathbf{u} + \mathbf{v}) = A\mathbf{u} + A\mathbf{v} = T_A(\mathbf{u}) + T_A(\mathbf{v})$$

y
$$T_A(c\mathbf{v}) = A(c\mathbf{v}) = c(A\mathbf{v}) = cT_A(\mathbf{v})$$

Por tanto, T_A es una transformación lineal.

Ejemplo 3.56

Sea $F: \mathbb{R}^2 \to \mathbb{R}^2$ la transformación que envía cada punto a su reflexión en el eje x. Demuestre que F es una transformación lineal.

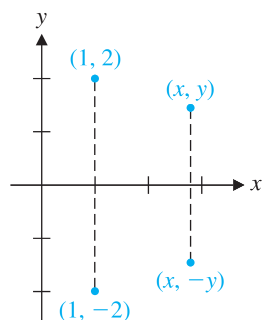


Figura 3.4
Reflexión en el eje x

Solución En la figura 3.4 es claro que F envía el punto (x, y) al punto $(x, -y)$. Por tanto, puede escribir

$$F\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}$$

Podría proceder para comprobar que F es lineal, como en el ejemplo 3.55 (¡esto incluso es más sencillo de comprobar!), pero es más rápido observar que

$$\begin{bmatrix} x \\ -y \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Por tanto, $F\begin{bmatrix} x \\ y \end{bmatrix} = A\begin{bmatrix} x \\ y \end{bmatrix}$, donde $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, de modo que F es una transformación matricial. Ahora se tiene, por el Teorema 3.30, que F es una transformación lineal.

Ejemplo 3.57

Sea $R: \mathbb{R}^2 \to \mathbb{R}^2$ la transformación que rota cada punto 90° en sentido contrario al de las manecillas del reloj en torno al origen. Demuestre que R es una transformación lineal.

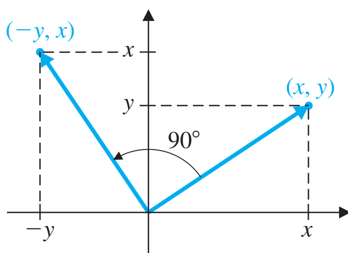


Figura 3.5
Una rotación de 90°

Solución Como muestra la figura 3.5, R envía el punto (x, y) al punto $(-y, x)$. Por tanto, se tiene

$$R\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -y \\ x \end{bmatrix} = x\begin{bmatrix} 0 \\ 1 \end{bmatrix} + y\begin{bmatrix} -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

En consecuencia, R es una transformación matricial y por lo tanto es lineal.

Observe que, si multiplica una matriz por vectores base estándar, obtiene las columnas de la matriz. Por ejemplo,

$$\begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ c \\ e \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} b \\ d \\ f \end{bmatrix}$$

Puede usar esta observación para demostrar que *toda* transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ surge como una transformación matricial.

Teorema 3.31 Sea $T : \mathbb{R}^n \to \mathbb{R}^m$ una transformación lineal. Entonces T es una transformación matricial. De manera más específica, $T = T_A$, donde A es la matriz de $m \times n$

$$A = [T(\mathbf{e}_1) \vdots T(\mathbf{e}_2) \vdots \cdots \vdots T(\mathbf{e}_n)]$$

Demostración Sean $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ los vectores base estándar en $\mathbb{R}^n$ y sea $\mathbf{x}$ un vector en $\mathbb{R}^n$. Puede escribir $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$ (donde las x_i son los componentes de $\mathbf{x}$). También se sabe que $T(\mathbf{e}_1), T(\mathbf{e}_2), \ldots, T(\mathbf{e}_n)$ son vectores (columna) en $\mathbb{R}^m$. Sea $A = [T(\mathbf{e}_1) \vdots T(\mathbf{e}_2) \vdots \cdots \vdots T(\mathbf{e}_n)]$ la matriz de $m \times n$ con dichos vectores como sus columnas. Entonces

$$\begin{aligned} T(\mathbf{x}) &= T(x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n) \\ &= x_1T(\mathbf{e}_1) + x_2T(\mathbf{e}_2) + \cdots + x_nT(\mathbf{e}_n) \\ &= [T(\mathbf{e}_1) \vdots T(\mathbf{e}_2) \vdots \cdots \vdots T(\mathbf{e}_n)] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = A\mathbf{x} \end{aligned}$$

como se requiere.

La matriz A en el Teorema 3.31 se llama ***matriz estándar de la transformación lineal*** T.

Ejemplo 3.58 Demuestre que una rotación en torno al origen, a través de un ángulo θ, define una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ y encuentre su matriz estándar.

Solución Sea R_θ la rotación. Se dará un argumento geométrico para establecer el hecho de que R_θ es lineal. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en $\mathbb{R}^2$. Si no son paralelos, entonces la figura 3.6(a) muestra la regla del paralelogramo que determina $\mathbf{u} + \mathbf{v}$. Si ahora se aplica R_θ, todo el paralelogramo rota un ángulo θ, como se muestra en la figura 3.6(b). Pero la diagonal de este paralelogramo debe ser $R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$, de nuevo, por la regla del paralelogramo. Por tanto, $R_\theta(\mathbf{u} + \mathbf{v}) = R_\theta(\mathbf{u}) + R_\theta(\mathbf{v})$. (¿Qué ocurre si $\mathbf{u}$ y $\mathbf{v}$ son paralelos?)

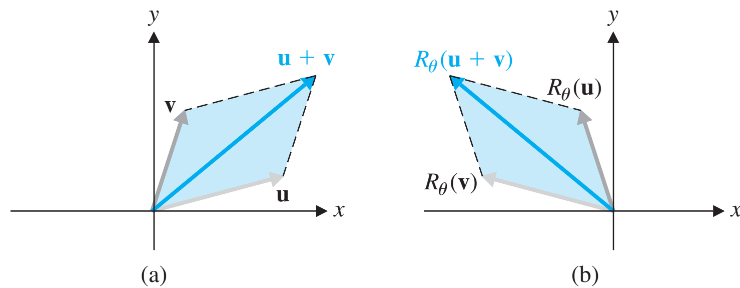


Figura 3.6

De igual modo, si se aplica R_θ a $\mathbf{v}$ y $c\mathbf{v}$, se obtiene $R_\theta(\mathbf{v})$ and $R_\theta(c\mathbf{v})$, como se muestra en la figura 3.7. Pero, dado que la rotación no afecta las longitudes, entonces se debe tener $R_\theta(c\mathbf{v}) = cR_\theta(\mathbf{v})$, como se requiere. (Dibuje diagramas para el caso $0 < c < 1$, $-1 < c < 0$ y $c < -1$.)

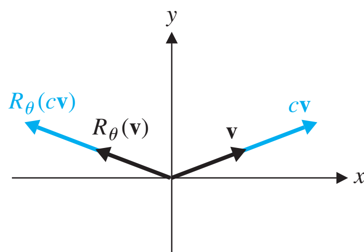


Figura 3.7

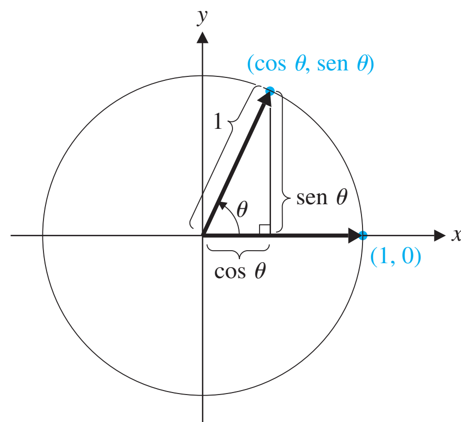


Figura 3.8
$R_\theta(\mathbf{e}_1)$

Por tanto, R_θ es una transformación lineal. De acuerdo con el Teorema 3.31, se puede encontrar su matriz al determinar su efecto sobre los vectores base estándar $\mathbf{e}_1$ y $\mathbf{e}_2$ de $\mathbb{R}^2$.

Ahora, como muestra la figura 3.8, $R_\theta\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}\cos\theta\\ \operatorname{sen}\theta\end{bmatrix}$.

Puede encontrar $R_\theta\begin{bmatrix}0\\1\end{bmatrix}$ de igual forma, pero es más rápido observar que $R_\theta\begin{bmatrix}0\\1\end{bmatrix}$ debe ser perpendicular (en sentido contrario al de las manecillas del reloj) a $R_\theta\begin{bmatrix}1\\0\end{bmatrix}$ y por tanto, por el ejemplo 3.57, $R_\theta\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}-\operatorname{sen}\theta\\ \cos\theta\end{bmatrix}$ (figura 3.9).

En consecuencia, la matriz estándar de R_θ es $\begin{bmatrix}\cos\theta & -\operatorname{sen}\theta\\ \operatorname{sen}\theta & \cos\theta\end{bmatrix}$.

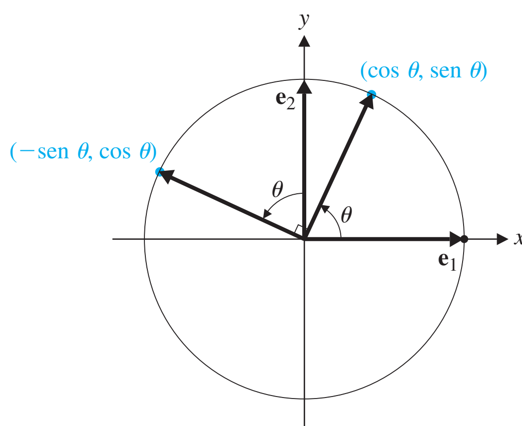


Figura 3.9
$R_\theta(\mathbf{e}_2)$

Ahora puede usar el resultado del ejemplo 3.58 para calcular el efecto de cualquier rotación. Por ejemplo, suponga que quiere rotar 60° el punto $(2, -1)$ en torno al origen. (La convención es que un ángulo positivo corresponde a una rotación en sentido con-

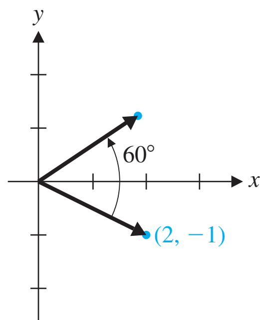


Figura 3.10
Una rotación de 60°

trario al de las manecillas del reloj, mientras que un ángulo negativo es en el sentido de las manecillas del reloj.) Dado que $\cos 60° = 1/2$ y $\operatorname{sen} 60° = \sqrt{3}/2$, se calcula

$$R_{60}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} \cos 60° & -\operatorname{sen} 60° \\ \operatorname{sen} 60° & \cos 60° \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$
$$= \begin{bmatrix} (2+\sqrt{3})/2 \\ (2\sqrt{3}-1)/2 \end{bmatrix}$$

Por tanto, la imagen del punto $(2, -1)$ bajo esta rotación es el punto $((2+\sqrt{3})/2, (2\sqrt{3}-1)/2) \approx (1.87, 1.23)$, como se muestra en la figura 3.10.

Ejemplo 3.59

(a) Demuestre que la transformación $P : \mathbb{R}^2 \to \mathbb{R}^2$ que proyecta un punto sobre el eje x es una transformación lineal y encuentre su matriz estándar.
(b) De manera más general, si ℓ es una recta que pasa por el origen en $\mathbb{R}^2$, demuestre que la transformación $P_\ell : \mathbb{R}^2 \to \mathbb{R}^2$ que proyecta un punto sobre ℓ es una transformación lineal y encuentre su matriz estándar.

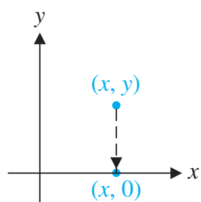


Figura 3.11
Una proyección

Solución (a) Como muestra la figura 3.11, P envía el punto (x, y) al punto $(x, 0)$. Por tanto,

$$P\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

Se tiene que P es una transformación matricial (y en consecuencia una transformación lineal) con matriz estándar $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

(b) Sea la recta ℓ que tiene vector director $\mathbf{d}$ y sea $\mathbf{v}$ un vector arbitrario. Entonces P_ℓ está dado por $\operatorname{proy}_{\mathbf{d}}(\mathbf{v})$, la proyección de $\mathbf{v}$ sobre $\mathbf{d}$, que de la sección 1.2 recordará tiene la fórmula

$$\operatorname{proy}_{\mathbf{d}}(\mathbf{v}) = \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d}$$

Por tanto, para demostrar que P_ℓ es lineal, se procede del modo siguiente:

$$\begin{aligned} P_\ell(\mathbf{u}+\mathbf{v}) &= \left(\frac{\mathbf{d}\cdot(\mathbf{u}+\mathbf{v})}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}+\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}}{\mathbf{d}\cdot\mathbf{d}} + \frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} \\ &= \left(\frac{\mathbf{d}\cdot\mathbf{u}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} + \left(\frac{\mathbf{d}\cdot\mathbf{v}}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = P_\ell(\mathbf{u}) + P_\ell(\mathbf{v}) \end{aligned}$$

De igual modo, $P_\ell(c\mathbf{v}) = cP_\ell(\mathbf{v})$ para cualquier escalar c (ejercicio 52). En consecuencia, P_ℓ es una transformación lineal.

Para encontrar la matriz estándar de P_ℓ, aplique el Teorema 3.31. Si se hace $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$, entonces

$$P_\ell(\mathbf{e}_1) = \left(\frac{\mathbf{d}\cdot\mathbf{e}_1}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = \frac{d_1}{d_1^2 + d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 \\ d_1 d_2 \end{bmatrix}$$

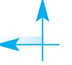
y

$$P_\ell(\mathbf{e}_2) = \left(\frac{\mathbf{d}\cdot\mathbf{e}_2}{\mathbf{d}\cdot\mathbf{d}}\right)\mathbf{d} = \frac{d_2}{d_1^2 + d_2^2}\begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1 d_2 \\ d_2^2 \end{bmatrix}$$

Por tanto, la matriz estándar de la proyección es

$$A = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 & d_1 d_2 \\ d_1 d_2 & d_2^2 \end{bmatrix} = \begin{bmatrix} d_1^2/(d_1^2 + d_2^2) & d_1 d_2/(d_1^2 + d_2^2) \\ d_1 d_2/(d_1^2 + d_2^2) & d_2^2/(d_1^2 + d_2^2) \end{bmatrix}$$

Como comprobación, note que en el inciso (a) podría tomar $\mathbf{d} = \mathbf{e}_1$ como un vector director para el eje x. Por tanto, $d_1 = 1$ y $d_2 = 0$, y se obtiene $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, como antes.



Nuevas transformaciones lineales a partir de otras anteriores

Si $T : \mathbb{R}^m \to \mathbb{R}^n$ y $S : \mathbb{R}^n \to \mathbb{R}^p$ son transformaciones lineales, entonces puede seguir T por S para formar la ***composición*** de las dos transformaciones, denotada $S \circ T$. Note que, para que $S \circ T$ tenga sentido, el codominio de T y el dominio de S deben coincidir (en este caso, ambos son $\mathbb{R}^n$) y la transformación compuesta resultante $S \circ T$ va del dominio de T al codominio de S (en este caso, $S \circ T : \mathbb{R}^m \to \mathbb{R}^p$). La figura 3.12 muestra de manera esquemática cómo funciona esta composición. La definición formal de composición de transformaciones se toma directamente de esta figura y es igual que la correspondiente definición de composición de funciones ordinarias:

$$(S \circ T)(\mathbf{v}) = S(T(\mathbf{v}))$$

Desde luego, le gustaría que $S \circ T$ también fuese una transformación lineal, y felizmente se encuentra que lo es. Esto se puede demostrar al mostrar que $S \circ T$ satisface la definición de transformación lineal (lo que se hará en el capítulo 6), pero, dado que por el momento se supone que las transformaciones lineales y las transformaciones matriciales son la misma cosa, es suficiente demostrar que $S \circ T$ es una transformación matricial. Se usará la notación $[T]$ para la matriz estándar de una transformación lineal T.

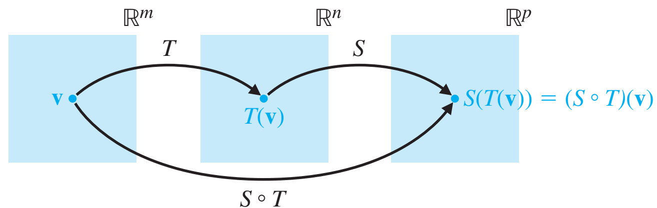


Figura 3.12
La composición de transformaciones

Teorema 3.32 Sean $T: \mathbb{R}^m \to \mathbb{R}^n$ y $S: \mathbb{R}^n \to \mathbb{R}^p$ transformaciones lineales. Entonces, $S \circ T: \mathbb{R}^m \to \mathbb{R}^p$ es una transformación lineal. Más aún, sus matrices estándar están relacionadas mediante

$$[S \circ T] = [S][T]$$

Demostración Sean $[S] = A$ y $[T] = B$. (Note que A es de $p \times n$ y B es de $n \times m$.) Si $\mathbf{v}$ es un vector en $\mathbb{R}^m$, entonces simplemente se calcula

$$(S \circ T)(\mathbf{v}) = S(T(\mathbf{v})) = S(B\mathbf{v}) = A(B\mathbf{v}) = (AB)\mathbf{v}$$



(Note aquí que las dimensiones de A y B garantizan que el producto AB tenga sentido.) Por tanto, se ve que el efecto de $S \circ T$ es multiplicar vectores por AB, de donde se tiene inmediatamente que $S \circ T$ es una transformación matricial (por tanto, lineal) con $[S \circ T] = [S][T]$.

¿No es un gran resultado? Dicho con palabras: "la matriz de la composición es el producto de la matrices". ¡Vaya fórmula adorable!

Ejemplo 3.60 Considere la transformación lineal $T: \mathbb{R}^2 \to \mathbb{R}^3$ del ejemplo 3.55, definida por

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 - x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

y la transformación lineal $S: \mathbb{R}^3 \to \mathbb{R}^4$ definida por

$$S\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 2y_1 + y_3 \\ 3y_2 - y_3 \\ y_1 - y_2 \\ y_1 + y_2 + y_3 \end{bmatrix}$$

Encuentre $S \circ T: \mathbb{R}^2 \to \mathbb{R}^4$.

Solución Se ve que las matrices estándar son

$$[S] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{y} \quad [T] = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix}$$

de modo que el Teorema 3.32 produce

$$[S \circ T] = [S][T] = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & -1 \\ 1 & -1 & 0 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix}$$

Se tiene que

$$(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 3 & -7 \\ -1 & 1 \\ 6 & 3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 5x_1 + 4x_2 \\ 3x_1 - 7x_2 \\ -x_1 + x_2 \\ 6x_1 + 3x_2 \end{bmatrix}$$

(En el ejercicio 29 se le pedirá comprobar este resultado al hacer

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 2x_1 - x_2 \\ 3x_1 + 4x_2 \end{bmatrix}$$

y sustituir estos valores en la definición de S, y de este modo calcular directamente $(S \circ T)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.)

Ejemplo 3.61 Encuentre la matriz estándar de la transformación que primero rota un punto 90° en sentido contrario al de las manecillas del reloj en torno al origen y luego refleja el resultado en el eje x.

Solución La rotación R y la reflexión F se discutieron en los ejemplos 3.57 y 3.56, respectivamente, de donde se encuentra que sus matrices estándar son $[R] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ y $[F] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Se tiene que la composición $F \circ R$ tiene por su matriz

$$[F \circ R] = [F][R] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

(Compruebe que este resultado es correcto al considerar el efecto de $F \circ R$ sobre los vectores base estándar $\mathbf{e}_1$ y $\mathbf{e}_2$. Note la importancia del *orden* de las transformaciones: R se realiza antes que F, pero se escribe $F \circ R$. En este caso, $R \circ F$ también tiene sentido. ¿$R \circ F = F \circ R$?)

Inversas de transformaciones lineales

Considere el efecto de una rotación de 90° en sentido contrario al de las manecillas del reloj en torno al origen, seguida por una rotación de 90° en el sentido de las manecillas del reloj. Claramente, esto deja todo punto en $\mathbb{R}^2$ sin cambios. Si estas transformaciones se denotan R_{90} y R_{-90} (recuerde que un ángulo negativo corresponde a dirección contraria a la de las manecillas del reloj), entonces puede expresar esto como $(R_{90} \circ R_{-90})(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{R}^2$. Note que, en este caso, si realiza las transformaciones en el otro orden, obtiene el mismo resultado final: $(R_{-90} \circ R_{90})(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{R}^2$.

Por tanto, $R_{90} \circ R_{-90}$ (y también $R_{-90} \circ R_{90}$) es una transformación lineal que deja todo vector en $\mathbb{R}^2$ sin cambio. Tal transformación se conoce como ***transformación identidad***. En general, se tiene una de tales transformaciones para todo $\mathbb{R}^n$; a saber, $I : \mathbb{R}^n \to \mathbb{R}^n$ tal que $I(\mathbf{v}) = \mathbf{v}$ para todo $\mathbf{v}$ en $\mathbb{R}^n$. (Si es importante seguir la pista de la dimensión del espacio, puede escribir I_n por claridad.)

De este modo, con esta notación, se tiene $R_{90} \circ R_{-90} = I = R_{-90} \circ R_{90}$. Un par de transformaciones que se relacionan mutuamente de esta forma se llaman ***transformaciones inversas***.

Definición Sean S y T transformaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^n$. Entonces S y T son ***transformaciones inversas*** si $S \circ T = I_n$ y $T \circ S = I_n$.

Comentario Dado que esta definición es simétrica con respecto a S y T, se dirá que, cuando ocurre esta situación, S es el inverso de T y T es el inverso de S. Más aún, se dirá que S y T son ***invertibles***.



En términos de matrices, se ve inmediatamente que si S y T son transformaciones inversas, entonces $[S][T] = [S \circ T] = [I] = I$, donde la última I es la *matriz* identidad. (¿Por qué la matriz estándar de la transformación identidad es la matriz identidad?) También se debe tener $[T][S] = [T \circ S] = [I] = I$. Esto demuestra que $[S]$ y $[T]$ son matrices inversas. Demuestra algo más: si una transformación lineal T es invertible, entonces su matriz estándar $[T]$ debe ser invertible, y dado que los inversos matriciales son únicos, esto significa que la inversa de T también es única. Por tanto, puede usar sin ambigüedad la notación T^{-1} para referirse a *la* inversa de T. Por tanto, puede reescribir las ecuaciones anteriores como $[T][T^{-1}] = I = [T^{-1}][T]$, lo que muestra que la matriz de T^{-1} es la matriz inversa de $[T]$. Se acaba de demostrar el siguiente teorema.

Teorema 3.33 Sea $T: \mathbb{R}^n \to \mathbb{R}^n$ una transformación lineal invertible. Entonces su matriz estándar $[T]$ es una matriz invertible y

$$[T^{-1}] = [T]^{-1}$$

Comentario También enuncie esto con palabras: "la matriz de la inversa es la inversa de la matriz". ¡Fabuloso!

Ejemplo 3.62 Encuentre la matriz estándar de una rotación de 60° en el sentido de las manecillas del reloj en torno al origen en $\mathbb{R}^2$.



Solución Anteriormente se calculó la matriz de una rotación de 60° en sentido contrario al de las manecillas del reloj en torno al origen como

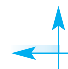
$$[R_{60}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$$

Dado que una rotación de 60° en el sentido de las manecillas del reloj es el inverso de una rotación de 60° en sentido contrario al de las manecillas del reloj, puede aplicar el Teorema 3.33 para obtener

$$[R_{-60}] = [(R_{60})^{-1}] = \begin{bmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}^{-1} = \begin{bmatrix} 1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & 1/2 \end{bmatrix}$$



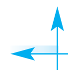
(Compruebe el cálculo de la inversa matricial. La forma más rápida es usar la abreviatura 2×2 del Teorema 3.8. También compruebe que la matriz resultante tiene el efecto correcto sobre la base estándar en $\mathbb{R}^2$ al dibujar un diagrama.)



Ejemplo 3.63 Determine si la proyección sobre el eje x es una transformación invertible y, si lo es, encuentre su inversa.



Solución La matriz estándar de esta proyección P es $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, que no es invertible pues su determinante es 0. Por tanto, P no es invertible tampoco.



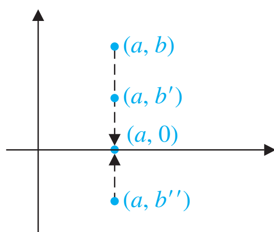


Figura 3.13
Las proyecciones no son invertibles

Comentario La figura 3.13 da alguna idea de por qué P en el ejemplo 3.63 no es invertible. La proyección "colapsa" $\mathbb{R}^2$ sobre el eje x. Para que P sea invertible, habría que tener una forma de "deshacerlo" para recuperar el punto (a, b) con el que se comenzó. Sin embargo, existe un número infinito de candidatos para la imagen de $(a, 0)$ bajo tal "inverso" hipotético. ¿Cuál usaría? No puede decir simplemente que P^{-1} debe enviar $(a, 0)$ a (a, b), pues esto no puede ser una *definición* cuando no hay forma de saber cuál debe ser b. (Vea el ejercicio 42.)

Asociatividad

El Teorema 3.3(a) de la sección 3.2 enuncia la propiedad de asociatividad para la multiplicación matricial: $A(BC) = (AB)C$. (Si no trató de demostrarlo entonces, hágalo ahora. Incluso con todas las matrices restringidas de 2×2 , obtendrá cierto sentimiento de la complejidad notacional involucrada en una prueba "por elementos", que debe hacerlo apreciar la demostración que está a punto de darse.)

El planteamiento de la demostración es vía transformaciones lineales. Se vio que toda matriz A de $n \times n$ da lugar a una transformación lineal $T_A : \mathbb{R}^n \to \mathbb{R}^m$ por el contrario, toda transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ tiene una correspondiente matriz $[T]$ de $m \times n$. Las dos correspondencias están inversamente relacionadas; esto es, dada A, $[T_A] = A$, y dada T, $T_{[T]} = T$.

Sea $R = T_A$, $S = T_B$ y $T = T_C$. Entonces, por el Teorema 3.32,

$$A(BC) = (AB)C \quad \text{si y sólo si} \quad R \circ (S \circ T) = (R \circ S) \circ T$$



Ahora se demuestra la última identidad. Sea $\mathbf{x}$ en el dominio de T (y por tanto en el dominio tanto de $R \circ (S \circ T)$ como de $(R \circ S) \circ T$; ¿por qué?). Para demostrar que $R \circ (S \circ T) = (R \circ S) \circ T$, es suficiente demostrar que tienen el mismo efecto sobre $\mathbf{x}$. Mediante la aplicación repetida de la definición de composición, se tiene

$$\begin{aligned}(R \circ (S \circ T))(\mathbf{x}) &= R((S \circ T)(\mathbf{x})) \\ &= R(S(T(\mathbf{x}))) \\ &= (R \circ S)(T(\mathbf{x})) = ((R \circ S) \circ T)(\mathbf{x})\end{aligned}$$



como se requería. (Compruebe cuidadosamente cómo se usó cuatro veces la definición de composición.)

Esta sección sirvió como introducción a las transformaciones lineales. En el capítulo 6 se observarán con más detalle y más generalidad estas transformaciones. Los ejercicios que siguen también contienen algunas exploraciones adicionales de este importante concepto.

Ejercicios 3.6

1. Sea $T_A : \mathbb{R}^2 \to \mathbb{R}^2$ la transformación matricial correspondiente a $A = \begin{bmatrix} 2 & -1 \\ 3 & 4 \end{bmatrix}$. Encuentre $T_A(\mathbf{u})$ y $T_A(\mathbf{v})$, donde $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

2. Sea $T_A : \mathbb{R}^2 \to \mathbb{R}^3$ la transformación matricial correspondiente a $A = \begin{bmatrix} 3 & -1 \\ 1 & 2 \\ 1 & 4 \end{bmatrix}$. Encuentre $T_A(\mathbf{u})$ y $T_A(\mathbf{v})$, donde $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

En los ejercicios 3-6, pruebe que la transformación dada es una transformación lineal, usando la definición (o el Comentario que sigue al ejemplo 3.55).

3. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + y \\ x - y \end{bmatrix}$ **4.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ -x \\ 3x - 7y \end{bmatrix}$

5. $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ 2x + y - 3z \end{bmatrix}$ **6.** $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ x + y \\ x + y + z \end{bmatrix}$

En los ejercicios 7-10, proporcione un contraejemplo para demostrar que la transformación dada no es una transformación lineal.

7. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x^2 \end{bmatrix}$ **8.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} |x| \\ |y| \end{bmatrix}$

9. $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} xy \\ x + y \end{bmatrix}$ **10.** $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 1 \\ y - 1 \end{bmatrix}$

En los ejercicios 11-14, encuentre la matriz estándar de la transformación lineal en el ejercicio dado.

11. Ejercicio 3 **12.** Ejercicio 4

13. Ejercicio 5 **14.** Ejercicio 6

En los ejercicios 15-18, demuestre que la transformación dada de $\mathbb{R}^2$ a $\mathbb{R}^2$ es lineal, al demostrar que es una transformación matricial.

15. F refleja un vector en el eje y.

16. R rota un vector 45° en sentido contrario al de las manecillas del reloj en torno al origen.

17. D estira un vector por un factor de 2 en el componente x y un factor de 3 en el componente y.

18. P proyecta un vector sobre la recta $y = x$.

19. Los tres tipos de matrices elementales dan lugar a cinco tipos de matrices de 2×2 con uno de las formas siguientes:

$$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix} \text{ o } \begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix} \text{ o } \begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$$

Cada una de estas matrices elementales corresponde a una transformación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$. Realice dibujos para ilustrar el efecto de cada una sobre el cuadrado unitario con vértices en (0, 0), (1, 0), (0, 1) y (1, 1).

En los ejercicios 20-25, encuentre la matriz estándar de la transformación lineal dada de $\mathbb{R}^2$ a $\mathbb{R}^2$.

20. Rotación en sentido contrario al de las manecillas del reloj de 120° en torno al origen.

21. Rotación en el sentido de las manecillas del reloj 30° en torno al origen.

22. Proyección sobre la recta $y = 2x$

23. Proyección sobre la recta $y = -x$

24. Reflexión en la recta $y = x$

25. Reflexión en la recta $y = -x$

26. Sea ℓ una recta que pasa por el origen en $\mathbb{R}^2$, P_ℓ la transformación lineal que proyecta un vector sobre ℓ y F_ℓ la transformación que refleja un vector en ℓ.

(a) Dibuje diagramas para demostrar que F_ℓ es lineal.

(b) La figura 3.14 sugiere una forma de encontrar la matriz de F_ℓ usando el hecho de que las diagonales de un paralelogramo se bisectan mutuamente. Demuestre que $F_\ell(\mathbf{x}) = 2P_\ell(\mathbf{x}) - \mathbf{x}$ y use este resultado para demostrar que la matriz estándar de F_ℓ es

$$\frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 - d_2^2 & 2d_1d_2 \\ 2d_1d_2 & -d_1^2 + d_2^2 \end{bmatrix}$$

(donde el vector director de ℓ es $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$).

(c) Si el ángulo entre ℓ y el eje x positivo es θ, demuestre que la matriz de F_ℓ es

$$\begin{bmatrix} \cos 2\theta & \text{sen } 2\theta \\ \text{sen } 2\theta & -\cos 2\theta \end{bmatrix}$$

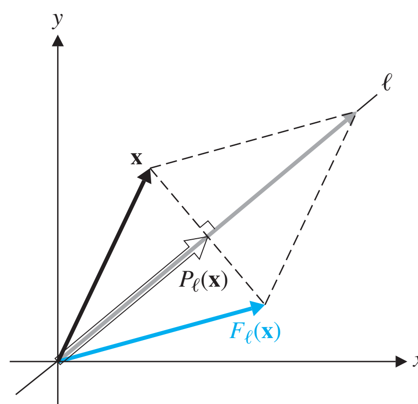


Figura 3.14

En los ejercicios 27 y 28, aplique el inciso (b) o (c) del ejercicio 26 para encontrar la matriz estándar de la transformación.

27. Reflexión en la recta $y = 2x$.

28. Reflexión en la recta $y = \sqrt{3}x$

29. Compruebe la fórmula para $S \circ T$ en el ejemplo 3.60 al realizar la sustitución directa sugerida.

En los ejercicios 30-35, verifique el Teorema 3.32 al encontrar la matriz de $S \circ T$ (a) mediante sustitución directa y (b) mediante multiplicación matricial de $[S][T]$.

30. $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 \\ x_1 + x_2 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 2y_1 \\ -y_2 \end{bmatrix}$

31. $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ -3x_1 + x_2 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 + 3y_2 \\ y_1 - y_2 \end{bmatrix}$

32. $T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_1 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 + 3y_2 \\ 2y_1 + y_2 \\ y_1 - y_2 \end{bmatrix}$

33. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 - x_3 \\ 2x_1 - x_2 + x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 4y_1 - 2y_2 \\ -y_1 + y_2 \end{bmatrix}$

34. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + 2x_2 \\ 2x_2 - x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 - y_2 \\ y_1 + y_2 \\ -y_1 + y_2 \end{bmatrix}$

35. $T\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ x_2 + x_3 \\ x_1 + x_3 \end{bmatrix}, S\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} y_1 - y_2 \\ y_2 - y_3 \\ -y_1 + y_3 \end{bmatrix}$

En los ejercicios 36-39, encuentre la matriz estándar de la transformación compuesta de $\mathbb{R}^2$ a $\mathbb{R}^2$.

36. Rotación en sentido contrario al de las manecillas del reloj de 60°, seguida por reflexión en la recta $y = x$

37. Reflexión en el eje y, seguida por rotación de 30° en sentido de las manecillas del reloj

38. Rotación de 45° en sentido de las manecillas del reloj, seguida por proyección sobre el eje y, seguida por rotación de 45° en el sentido de las manecillas del reloj

39. Reflexión en la recta $y = x$, seguida por rotación de 30° en sentido contrario al de las manecillas del reloj, seguida por reflexión en la recta $y = -x$

En los ejercicios 40-43, use matrices para probar los enunciados dados acerca de las transformaciones de $\mathbb{R}^2$ en $\mathbb{R}^2$.

40. Si R_θ denota una rotación (en torno al origen) en un ángulo θ, entonces $R_\alpha \circ R_\beta = R_{\alpha+\beta}$.

41. Si θ es el ángulo entre las rectas ℓ y m (que pasan por el origen), entonces $F_m \circ F_\ell = R_{+2\theta}$. (Vea el ejercicio 26.)

42. **(a)** Si P es una proyección, entonces $P \circ P = P$.
(b) La matriz de una proyección nunca puede ser invertible.

43. Si ℓ, m y n son tres rectas que pasan por el origen, entonces $F_n \circ F_m \circ F_\ell$ es también una reflexión que pasa por el origen.

44. Sea T una transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^2$ (o de $\mathbb{R}^3$ a $\mathbb{R}^3$). Demuestre que T mapea una recta en una recta o un punto. [*Sugerencia:* use la forma vectorial de la ecuación de una recta.]

45. Sea T una transformación lineal de $\mathbb{R}^2$ a $\mathbb{R}^2$ (o de $\mathbb{R}^3$ a $\mathbb{R}^3$). Pruebe que T mapea rectas paralelas a rectas paralelas, una sola recta, un par de puntos, o un solo punto.

En los ejercicios 46-51, sea ABCD el cuadrado con vértices $(-1, 1), (1, 1), (1, -1)$ *y* $(-1, -1)$. *Use los resultados en los ejercicios 44 y 45 para encontrar y dibujar la imagen de ABCD de acuerdo con la transformación dada.*

46. T en el ejercicio 3

47. D en el ejercicio 17

48. P en el ejercicio 18

49. La proyección en el ejercicio 22

50. T en el ejercicio 31

51. La transformación en el ejercicio 37

52. Demuestre que $P_\ell(c\mathbf{v}) = cP_\ell(\mathbf{v})$ para cualquier escalar c [Ejemplo 3.59(b)].

53. Demuestre que $T : \mathbb{R}^n \to \mathbb{R}^m$ es una transformación lineal si y sólo si

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)$$

para todo $\mathbf{v}_1, \mathbf{v}_2$ en $\mathbb{R}^n$ y escalares c_1, c_2.

54. Demuestre que (como se anotó al comienzo de esta sección) el rango de una transformación lineal $T : \mathbb{R}^n \to \mathbb{R}^m$ es el espacio columna de su matriz $[T]$.

55. Si A es una matriz invertible de 2×2, ¿qué afirma el teorema fundamental de las matrices invertibles acerca de la correspondiente transformación lineal T_A a la luz del ejercicio 19?

Robótica

En 1981, el transbordador espacial estadounidense *Columbia* despegó con un dispositivo llamado Shuttle Remote Manipulador System (SRMS, sistema manipulador remoto del transbordador). Este brazo robótico, conocido como Canadarm, resultó ser una herramienta vital en todas las misiones posteriores del transbordador espacial, pues proporcionó una manipulación firme, aunque precisa y delicada, de sus cargas (vea la figura 3.15).

Canadarm se ha usado para poner satélites en su órbita adecuada y para recuperar los que funcionan mal para repararlos, y también ha realizado reparaciones cruciales al mismo transbordador. Notablemente, el brazo robótico fue útil en la reparación exitosa del *Telescopio Espacial Hubble*. Desde 1998, Canadarm tiene un papel importante en el ensamblado de la *Estación Espacial Internacional*.



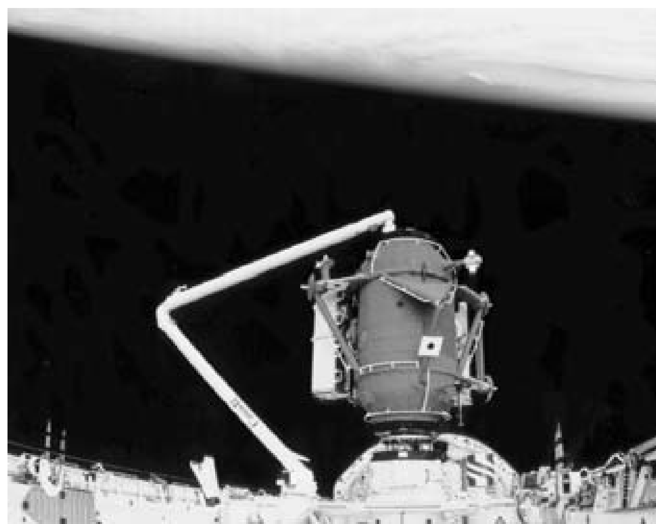


Figura 3.15
Canadarm

Un brazo robótico consiste de una serie de *vástagos* de longitud fija conectados en *articulaciones* donde pueden rotar. En consecuencia, cada vástago rota en el espacio o (mediante el efecto de los otros vástagos) se traslada paralelo a sí mismo, o se mueve mediante una combinación (composición) de rotaciones y traslaciones. Antes de poder diseñar un modelo matemático para un brazo robótico, es necesario entender cómo funcionan las rotaciones y traslaciones en composición. Para simplificar las cosas, se supondrá que el brazo está en $\mathbb{R}^2$.

En la sección 3.6 se vio que la matriz de una rotación R en torno al origen en un ángulo θ es una transformación lineal con matriz $\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ [figura 3.16(a)]. Si $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$, entonces una ***traslación a lo largo de*** $\mathbf{v}$ es la transformación

$$T(\mathbf{x}) = \mathbf{x} + \mathbf{v} \quad \text{o de manera equivalente,} \quad T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + a \\ y + b \end{bmatrix}$$

[figura 3.16(b)].

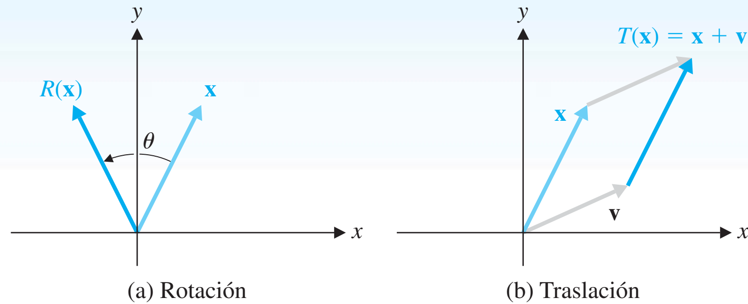


Figura 3.16

Por desgracia, la traslación no es una transformación lineal, porque $T(\mathbf{0}) \neq \mathbf{0}$. Sin embargo, hay un truco que dará la vuelta a este problema. Puede representar el vector $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ como el vector $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$ en $\mathbb{R}^3$. A esto se le llama representar $\mathbf{x}$ en ***coordenadas homogéneas***. Entonces la multiplicación matricial

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x + a \\ y + b \\ 1 \end{bmatrix}$$

representa al vector trasladado $T(\mathbf{x})$ en coordenadas homogéneas.

También puede tratar las rotaciones en coordenadas homogéneas. La multiplicación matricial

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x\cos\theta - y\operatorname{sen}\theta \\ x\operatorname{sen}\theta + y\cos\theta \\ 1 \end{bmatrix}$$

representa el vector rotado $R(\mathbf{x})$ en coordenadas homogéneas. La composición $T \circ R$ que produce la rotación R seguida por la traslación T ahora se representa mediante el producto

$$\begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & 0 \\ \operatorname{sen}\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta & a \\ \operatorname{sen}\theta & \cos\theta & b \\ 0 & 0 & 1 \end{bmatrix}$$

(Note que $R \circ T \neq T \circ R$.)

Para modelar un brazo robótico, proporcione a cada vástago su propio sistema coordenado (llamado *marco*) y examine cómo se mueve un vástago en relación con aque-

llos con los que está directamente conectado. Para ser específico, sean X_i y Y_i los ejes coordenados del vástago $\mathcal{A}_i$ con el eje X_i alineado con el vástago. La longitud de $\mathcal{A}_i$ se denota mediante a_i y el ángulo entre X_i y X_{i-1} se denota con θ_i. La articulación entre $\mathcal{A}_i$ y $\mathcal{A}_{i-1}$ está en el punto $(0, 0)$ con respecto a $\mathcal{A}_i$ y $(a_{i-1}, 0)$ relativo a $\mathcal{A}_{i-1}$. Por tanto, en relación con $\mathcal{A}_{i-1}$, el sistema coordenado para $\mathcal{A}_i$ rotó un ángulo de θ_i y luego se trasladó a lo largo de $\begin{bmatrix} a_{i-1} \\ 0 \end{bmatrix}$ (figura 3.17). Esta transformación se representa en coordenadas homogéneas mediante la matriz

$$T_i = \begin{bmatrix} \cos\theta_i & -\operatorname{sen}\theta_i & a_{i-1} \\ \operatorname{sen}\theta_i & \cos\theta_i & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

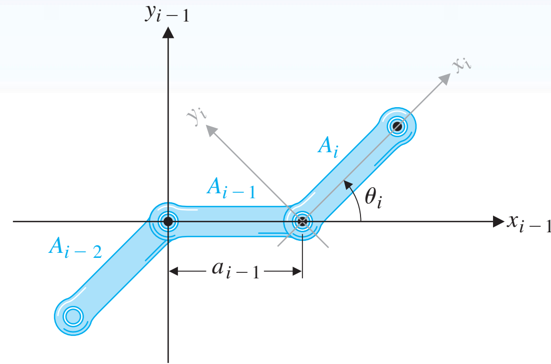


Figura 3.17

Para dar un ejemplo específico, considere la figura 3.18(a). En ella se muestra un brazo con tres vástagos en el que $\mathcal{A}_1$ está en su posición inicial y cada uno de los otros dos vástagos rotó 45° desde el vástago previo. La longitud de cada vástago se considerará en 2 unidades. La figura 3.18(b) muestra $\mathcal{A}_3$ en su marco inicial. La transformación

$$T_3 = \begin{bmatrix} \cos 45 & -\operatorname{sen} 45 & 2 \\ \operatorname{sen} 45 & \cos 45 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 2 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

causa una rotación de 45° y luego una traslación de 2 unidades. Como se muestra en 3.18(c), esto coloca a $\mathcal{A}_3$ en su posición adecuada en relación con el marco de $\mathcal{A}_2$. A continuación, la transformación

$$T_2 = \begin{bmatrix} \cos 45 & -\operatorname{sen} 45 & 2 \\ \operatorname{sen} 45 & \cos 45 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 2 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

se aplica al resultado anterior. Esto coloca tanto a $\mathcal{A}_3$ como a $\mathcal{A}_2$ en su posición correcta en relación con $\mathcal{A}_1$, como se muestra en la figura 3.18(d). Por lo general, al resultado previo se aplicaría una tercera transformación T_1 (una rotación), pero en este caso, T_1 es la transformación identidad porque $\mathcal{A}_1$ permanece en su posición inicial.

Usualmente se quiere conocer las coordenadas del extremo (la "mano") del brazo robótico, dados los parámetros de longitud y ángulo; esto se conoce como *forward kinematics*, cinemática hacia adelante. Al continuar con la secuencia de cálculos anterior y

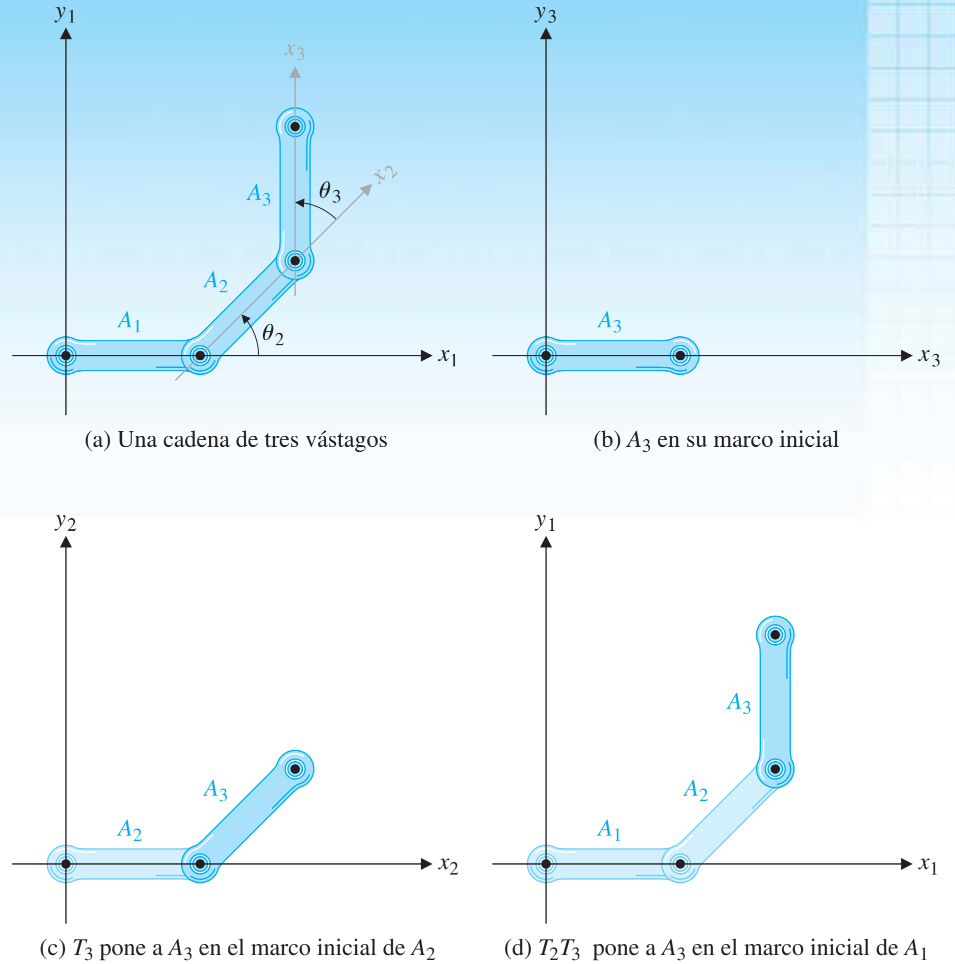


Figure 3.18

consultar la figura 3.18, se ve que es necesario determinar dónde termina el punto $(2, 0)$ después de aplicar T_3 y T_2. Por tanto, la mano del brazo está en

$$T_2T_3\begin{bmatrix}2\\0\\1\end{bmatrix} = \begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2} & 2\\ 1/\sqrt{2} & 1/\sqrt{2} & 0\\ 0 & 0 & 1\end{bmatrix}^2\begin{bmatrix}2\\0\\1\end{bmatrix} = \begin{bmatrix}0 & -1 & 2+\sqrt{2}\\ 1 & 0 & \sqrt{2}\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}2\\0\\1\end{bmatrix}$$
$$= \begin{bmatrix}2+\sqrt{2}\\ 2+\sqrt{2}\\ 1\end{bmatrix}$$

que representa el punto $(2+\sqrt{2}, 2+\sqrt{2})$ en coordenadas homogéneas. Fácilmente se comprueba, a partir de la figura 3.18(a), que esto es correcto.

Los métodos empleados en este ejemplo se generalizan a brazos robóticos en tres dimensiones, aunque en $\mathbb{R}^3$ existen más grados de libertad y por tanto más variables. El método de coordenadas homogéneas también es útil en otras aplicaciones, notablemente en gráficos de computadora.

3.7 Aplicaciones

Cadenas de Markov

Un equipo de investigación de mercado realiza un estudio controlado para determinar qué dentífricos prefieren las personas. La muestra consiste de 200 personas, y a cada una se le pide probar dos marcas de dentífrico durante un periodo de varios meses. Con base en las respuestas de la encuesta, el equipo de investigación compila las siguientes estadísticas acerca de las preferencias de dentífrico.

De quienes usan la marca A en cualquier mes, 70% siguen usándola el mes siguiente, mientras que 30% cambia a la marca B; de quienes usan la marca B en cualquier mes, 80% siguen usándolo el mes siguiente, mientras que 20% cambian a la marca A. Estos hallazgos se resumen en la figura 3.19, en la que los porcentajes se convirtieron en decimales; se pensará en ellos como probabilidades.

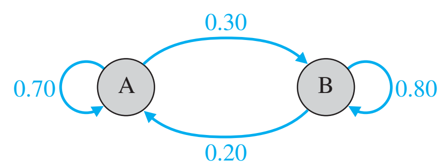


Figura 3.19

Andrei A. Markov (1856–1922) fue un matemático ruso que estudió y más tarde impartió clases en la Universidad de San Petersburgo. Estuvo interesado en teoría de números, análisis y teoría de las fracciones continuas, un campo recientemente desarrollado que Markov aplicó a la teoría de probabilidad. Markov también estuvo interesado en la poesía y uno de los usos que dio a las cadenas de Markov fue el análisis de patrones en poemas y otros textos literarios.

La figura 3.19 es un ejemplo simple de una ***cadena de Markov*** (finita). Representa un proceso evolutivo que consiste de un número finito de *estados*. En cada paso o punto en el tiempo, el proceso puede estar en cualquiera de los estados; en el paso siguiente, el proceso puede permanecer en su estado presente o cambiar a uno de los otros estados. El estado hacia donde avanza el proceso en el siguiente paso y la probabilidad de hacerlo depende *solamente* del estado presente y no de la historia pasada del proceso. Estas probabilidades se conocen como *probabilidades de transición* y se suponen constantes (esto es: la probabilidad de moverse del estado i al estado j siempre es la misma).

Ejemplo 3.64

En el estudio de dentífricos descrito anteriormente, existen sólo dos estados (usar la marca A y usar la marca B) y las probabilidades de transición son las indicadas en la figura 3.19. Suponga que, cuando comienza el estudio, 120 personas usan la marca A y 80 personas usan la marca B. ¿Cuántas personas usarán cada marca 1 mes más tarde? ¿Y 2 meses más tarde?

Solución El número de usuarios de la marca A después de 1 mes será 70% de los que inicialmente usan la marca A (quienes siguen siendo leales a la marca A) más 20% de los usuarios de la marca B (quienes cambian de B a A):

$$0.70(120) + 0.20(80) = 100$$

De igual modo, el número de usuarios de la marca B después de 1 mes será una combinación de quienes cambian a la marca B y quienes siguen usándola:

$$0.30(120) + 0.80(80) = 100$$

Estas dos ecuaciones pueden resumirse en una sola ecuación matricial:

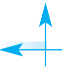
$$\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 120 \\ 80 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$$

Llame P a la matriz y marque los vectores $\mathbf{x}_0 = \begin{bmatrix} 120 \\ 80 \end{bmatrix}$ y $\mathbf{x}_1 = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$. (Note que los componentes de cada vector son el número de usuarios de las marcas A y B, en ese orden, después del número de meses indicado por el subíndice.) Por tanto, se tiene $\mathbf{x}_1 = P\mathbf{x}_0$.

Al extender la notación, sea $\mathbf{x}_k$ el vector cuyos componentes registran la distribución de usuarios de dentífrico después de k meses. Para determinar el número de usuarios de cada marca después de transcurridos 2 meses, simplemente aplique el mismo razonamiento comenzando con $\mathbf{x}_1$ en lugar de $\mathbf{x}_0$. Se obtiene

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \begin{bmatrix} 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 90 \\ 110 \end{bmatrix}$$

de donde se ve que ahora existen 90 usuarios de la marca A y 110 usuarios de la marca B.



Los vectores $\mathbf{x}_k$ en el ejemplo 3.64 se llaman ***vectores de estado*** de la cadena de Markov, y la matriz P se llama ***matriz de transición***. Acaba de ver que una cadena de Markov satisface la relación

$$\mathbf{x}_{k+1} = P\mathbf{x}_k \quad \text{para} \quad k = 0, 1, 2, \ldots$$

A partir de este resultado se tiene que es posible calcular un vector de estado arbitrario *de manera iterativa* una vez se conoce $\mathbf{x}_0$ y P. En otras palabras, una cadena de Markov está *completamente determinada* por sus probabilidades de transición y su estado inicial.

Comentarios

- Suponga que en el ejemplo 3.64 se quiere seguir la pista no de los números *reales* de usuarios de dentífrico, sino de los números *relativos* que usan cada marca. Podría convertir los datos en porcentajes o fracciones al dividir el número total de usuarios entre 200. Por tanto, comenzaría con

$$\mathbf{x}_0 = \begin{bmatrix} \frac{120}{200} \\ \frac{80}{200} \end{bmatrix} = \begin{bmatrix} 0.60 \\ 0.40 \end{bmatrix}$$

para reflejar el hecho de que, inicialmente, la división marca A-marca B es 60%–40%. Compruebe mediante cálculo directo que $P\mathbf{x}_0 = \begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}$, que entonces puede tomarse como $\mathbf{x}_1$ (en concordancia con la división 50-50 que se calculó anteriormente). Los vectores como este, con componentes no negativos que suman 1, se llaman ***vectores de probabilidad***.

- Observe cómo se ordenan las probabilidades de transición dentro de la matriz de transición P. Puede considerar a las columnas como etiquetadas con los estados *presente* y los renglones como etiquetados con los estados *siguiente*:

$$\begin{array}{cc} & \textit{Presente} \\ & \begin{array}{cc} \text{A} & \text{B} \end{array} \\ \textit{Siguiente} \begin{array}{c} \text{A} \\ \text{B} \end{array} & \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} \end{array}$$

La palabra *estocástico* se deriva del adjetivo griego *stokhastikos*, que significa "capaz de conjeturar" (o adivinar). Llegó a aplicarse a algo que está gobernado por las leyes de la probabilidad en el sentido de que la probabilidad realiza predicciones acerca de la posibilidad de que ocurran las cosas. En teoría de probabilidad, los "procesos estocásticos" forman una generalización de las cadenas de Markov.

Note también que las columnas de P son vectores de probabilidad; cualquier matriz cuadrada con esta propiedad se llama ***matriz estocástica***.

Puede darse cuenta de la naturaleza determinista de las cadenas de Markov en otra forma. Note que puede escribir

$$\mathbf{x}_2 = P\mathbf{x}_1 = P(P\mathbf{x}_0) = P^2\mathbf{x}_0$$

y, en general,

$$\mathbf{x}_k = P^k\mathbf{x}_0 \quad \text{para} \quad k = 0, 1, 2, \ldots$$

Esto conduce a examinar las potencias de una matriz de transición. En el ejemplo 3.64 se tiene

$$P^2 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix} = \begin{bmatrix} 0.55 & 0.30 \\ 0.45 & 0.70 \end{bmatrix}$$

¿Qué se hará con las entradas de esta matriz? Lo primero que debe observar es que p^2 es otra matriz estocástica, pues sus columnas suman 1. (Se le pedirá probar esto en el ejercicio 14.) ¿Podría ser que P^2 también sea una matriz de transición del mismo tipo? Considere una de sus entradas, por ejemplo, $(P^2)_{21} = 0.45$. Los tres diagramas de la figura 3.20 clarifican de dónde provienen estas entradas.

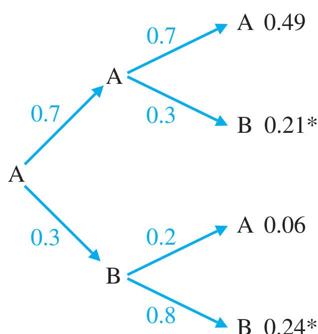


Figura 3.20

Existen cuatro posibles cambios de estado que pueden ocurrir durante 2 meses y corresponden a las cuatro ramas (o rutas) de longitud 2 en el árbol. Alguien que inicialmente use la marca A puede terminar usando la marca B dos meses después en dos formas diferentes (marcadas * en la figura): la persona puede seguir usando A después de 1 mes y luego cambiar a B (con probabilidad $0.7(0.3) = 0.21$), o la persona puede cambiar a B después de 1 mes y luego seguir con B (con probabilidad $0.3(0.8) = 0.24$). La suma de estas probabilidades produce una probabilidad global de 0.45. Observe que dichos cálculos son *exactamente* lo que se hizo cuando se calculó $(P^2)_{21}$.

Se tiene que $(P^2)_{21} = 0.45$ representa la probabilidad de moverse del estado 1 (marca A) al estado 2 (marca B) en dos transiciones. (Note que el orden de los subíndices es el *inverso* del que pudo suponer.) El argumento puede generalizarse para demostrar que

$(P^k)_{ij}$ es la probabilidad de moverse del estado j al estado i en k transiciones.

En el ejemplo 3.64, ¿qué sucederá con la distribución de usuarios de dentífrico a largo plazo? Trabaje con vectores de probabilidad como vectores de estado. Al continuar con los cálculos (redondeados a tres lugares decimales), se encuentra

$$\mathbf{x}_0 = \begin{bmatrix} 0.60 \\ 0.40 \end{bmatrix}, \mathbf{x}_1 = \begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix}, \mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 0.50 \\ 0.50 \end{bmatrix} = \begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix},$$

$$\mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 0.45 \\ 0.55 \end{bmatrix} = \begin{bmatrix} 0.425 \\ 0.575 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 0.412 \\ 0.588 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 0.406 \\ 0.594 \end{bmatrix},$$

$$\mathbf{x}_6 = \begin{bmatrix} 0.403 \\ 0.597 \end{bmatrix}, \mathbf{x}_7 = \begin{bmatrix} 0.402 \\ 0.598 \end{bmatrix}, \mathbf{x}_8 = \begin{bmatrix} 0.401 \\ 0.599 \end{bmatrix}, \mathbf{x}_9 = \begin{bmatrix} 0.400 \\ 0.600 \end{bmatrix}, \mathbf{x}_{10} = \begin{bmatrix} 0.400 \\ 0.600 \end{bmatrix},$$

etcétera. Parece que los vectores de estado tienden (o *convergen a*) el vector $\begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$, lo que implica que, a la larga, 40% de los usuarios de dentífrico en el estudio usarán la marca A y 60% usarán la marca B. De hecho, es fácil comprobar que, una vez alcanzada esta distribución, nunca cambiará. Simplemente calcule

$$\begin{bmatrix} 0.70 & 0.20 \\ 0.30 & 0.80 \end{bmatrix}\begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$$

Un vector de estado $\mathbf{x}$ con la propiedad de que $P\mathbf{x} = \mathbf{x}$ se llama ***vector de estado estacionario***. En el capítulo 4 se probará que toda cadena de Markov tiene un vector de estado estacionario único. Por ahora, acepte esto como un hecho y vea cómo puede encontrar tal vector sin hacer iteración alguna.

Comience por reescribir la ecuación matricial $P\mathbf{x} = \mathbf{x}$ como $P\mathbf{x} = I\mathbf{x}$, que a su vez puede reescribirse como $(I - P)\mathbf{x} = \mathbf{0}$. Ahora, esto es justo un sistema homogéneo de ecuaciones lineales con matriz de coeficientes $I - P$, de modo que la matriz aumentada es $[I - P|\mathbf{0}]$. En el ejemplo 3.64 se tiene

$$[I - P \mid \mathbf{0}] = \left[\begin{array}{cc|c} 1 - 0.70 & -0.20 & 0 \\ -0.30 & 1 - 0.80 & 0 \end{array}\right] = \left[\begin{array}{cc|c} 0.30 & -0.20 & 0 \\ -0.30 & 0.20 & 0 \end{array}\right]$$

que se reduce a

$$\left[\begin{array}{cc|c} 1 & -\frac{2}{3} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

De modo que, si el vector de estado estacionario es $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, entonces x_2 es una variable libre y la solución paramétrica es

$$x_1 = \tfrac{2}{3}t, \quad x_2 = t$$

Si se requiere que $\mathbf{x}$ sea un vector de probabilidad, entonces debe tener

$$1 = x_1 + x_2 = \tfrac{2}{3}t + t = \tfrac{5}{3}t$$

Por tanto, $x_2 = t = \frac{3}{5} = 0.6$ y $x_1 = \frac{2}{5} = 0.4$, de modo que $\mathbf{x} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$, en concordancia con los anteriores cálculos iterativos. (Si requiere que $\mathbf{x}$ contenga la distribución *real*, entonces en este ejemplo debe tener $x_1 + x_2 = 200$, de donde se tiene que $\mathbf{x} = \begin{bmatrix} 80 \\ 120 \end{bmatrix}$.)

Ejemplo 3.65

Un psicólogo coloca una rata en una jaula con tres compartimentos, como se muestra en la figura 3.21. La rata fue entrenada para seleccionar una puerta al azar siempre que suene una campana y moverse a través de ella hacia el siguiente compartimento.

(a) Si inicialmente la rata está en el compartimiento 1, ¿cuál es la probabilidad de que estará en el compartimiento 2 después de que la campana suene dos veces? ¿Tres veces?

(b) A largo plazo, ¿qué proporción de su tiempo pasará la rata en cada compartimiento?

Solución Sea $P = [p_{ij}]$ la matriz de transición para esta cadena de Markov. Entonces

$$p_{21} = p_{31} = \tfrac{1}{2},\; p_{12} = p_{13} = \tfrac{1}{3},\; p_{32} = p_{23} = \tfrac{2}{3} \;\text{ y }\; p_{11} = p_{22} = p_{33} = 0$$

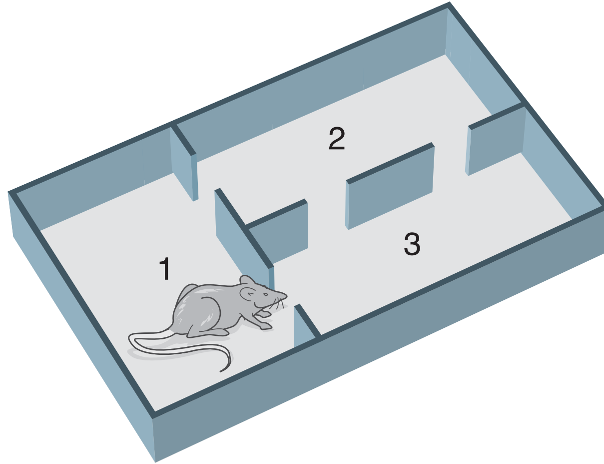


Figura 3.21



(¿Por qué? Recuerde que p_{ij} es la probabilidad de moverse de j a i.) Por tanto

$$P = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}$$

y el vector de estado inicial es

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

(a) Después de un campanazo, se tiene

$$\mathbf{x}_1 = P\mathbf{x}_0 = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0.5 \\ 0.5 \end{bmatrix}$$

Al continuar (redondeado a tres lugares decimales), se encuentra

$$\mathbf{x}_2 = P\mathbf{x}_1 = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}\begin{bmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} \approx \begin{bmatrix} 0.333 \\ 0.333 \\ 0.333 \end{bmatrix}$$

y

$$\mathbf{x}_3 = P\mathbf{x}_2 = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}\begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{2}{9} \\ \frac{7}{18} \\ \frac{7}{18} \end{bmatrix} \approx \begin{bmatrix} 0.222 \\ 0.389 \\ 0.389 \end{bmatrix}$$

Por tanto, después de dos campanazos, la probabilidad de que la rata esté en el compartimiento 2 es $\frac{1}{3} \approx 0.333$, y después de tres campanazos, la probabilidad de que la rata esté en el compartimiento 2 es $\frac{7}{18} \approx 0.389$. [Note que estas preguntas también podrían responderse al calcular $(P^2)_{21}$ y $(P^3)_{21}$.]

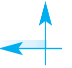
(b) Esta pregunta pide el vector de estado estacionario **x** como un vector de probabilidad. Como se vio anteriormente, **x** debe estar en el espacio nulo de $I - P$, de modo que se procede a resolver el sistema

$$[I - P \mid \mathbf{0}] = \left[\begin{array}{ccc|c} 1 & -\frac{1}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{2} & 1 & -\frac{2}{3} & 0 \\ -\frac{1}{2} & -\frac{2}{3} & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -\frac{2}{3} & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Por tanto, si $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, entonces $x_3 = t$ es libre y $x_1 = \frac{2}{3}t$, $x_2 = t$. Dado que **x** debe ser un vector de probabilidad, se necesita $1 = x_1 + x_2 + x_3 = \frac{8}{3}t$. En consecuencia, $t = \frac{3}{8}$ y

$$\mathbf{x} = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{8} \\ \frac{3}{8} \end{bmatrix}$$

lo cual dice que, a largo plazo, la rata pasa $\frac{1}{4}$ de su tiempo en el compartimiento 1 y $\frac{3}{8}$ de su tiempo en cada uno de los otros dos compartimientos.



Modelos económicos lineales

Ahora se visitarán nuevamente los modelos económicos que se encontraron por primera vez en la sección 2.4 y se reformularán en términos de matrices. El ejemplo 2.33 ilustró el modelo cerrado de Leontief. El sistema de ecuaciones que debió resolverse fue

$$\begin{aligned} \tfrac{1}{4}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{2}x_3 &= x_1 \\ \tfrac{1}{4}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{4}x_3 &= x_2 \\ \tfrac{1}{2}x_1 + \tfrac{1}{3}x_2 + \tfrac{1}{4}x_3 &= x_3 \end{aligned}$$

En forma matricial, esta es la ecuación $E\mathbf{x} = \mathbf{x}$, donde

$$E = \begin{bmatrix} 1/4 & 1/3 & 1/2 \\ 1/4 & 1/3 & 1/4 \\ 1/2 & 1/3 & 1/4 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

La matriz E se llama ***matriz de intercambio*** y el vector **x** se llama ***vector de precio***. En general, si $E = [e_{ij}]$, entonces e_{ij} representa la fracción (o porcentaje) de output de la industria j que consume la industria i, y x_i es el precio que cobra la industria i por su output.

En una economía cerrada, la suma de cada columna de E es 1. Dado que las entradas de E también son no negativas, E es una matriz estocástica y el problema de encontrar una solución a la ecuación

$$E\mathbf{x} = \mathbf{x} \tag{1}$$

¡es precisamente el mismo que el problema de encontrar el vector de estado estacionario de una cadena de Markov! Por tanto, para encontrar un vector precio **x** que satisfaga $E\mathbf{x} = \mathbf{x}$, se resuelve la ecuación homogénea equivalente $(I - E)\mathbf{x} = 0$. Siempre existirá un número infinito de soluciones; se busca una solución donde los precios sean todos no negativos y al menos un precio sea positivo.

El modelo abierto de Leontief es más interesante. En el ejemplo 2.34 fue necesario resolver el sistema

$$\begin{aligned} x_1 &= 0.2x_1 + 0.5x_2 + 0.1x_3 + 10 \\ x_2 &= 0.4x_1 + 0.2x_2 + 0.2x_3 + 10 \\ x_3 &= 0.1x_1 + 0.3x_2 + 0.3x_3 + 30 \end{aligned}$$

En forma matricial, se tiene

$$\mathbf{x} = C\mathbf{x} + \mathbf{d} \quad o \quad (I - C)\mathbf{x} = \mathbf{d} \tag{2}$$

donde

$$C = \begin{bmatrix} 0.2 & 0.5 & 0.1 \\ 0.4 & 0.2 & 0.2 \\ 0.1 & 0.3 & 0.3 \end{bmatrix}, \ \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \ \mathbf{d} = \begin{bmatrix} 10 \\ 10 \\ 30 \end{bmatrix}$$

La matriz C se llama ***matriz de consumo***, $\mathbf{x}$ es el ***vector de producción***, y $\mathbf{d}$ es el ***vector de demanda***. En general, si $C = [c_{ij}]$, $\mathbf{x} = [x_i]$, y $\mathbf{d} = [d_i]$, entonces c_{ij} representa el valor en dólares del output de la industria i que se necesita para producir un dólar de output de la industria j, x_i es el valor monetario (precio) del output de la industria i, y d_i es el valor monetario de la demanda externa para el output de la industria i. Una vez más, se está interesado en encontrar un vector de producción $\mathbf{x}$ con entradas no negativas tales que al menos una entrada sea positiva. A tal vector $\mathbf{x}$ se le llama ***solución factible***.

Ejemplo 3.66

Determine si existe una solución para modelo abierto de Leontief determinado por las siguientes matrices de consumo:

(a) $C = \begin{bmatrix} 1/4 & 1/3 \\ 1/2 & 1/3 \end{bmatrix}$ (b) $C = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 2/3 \end{bmatrix}$

Solución (a) Se tiene

$$I - C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1/4 & 1/3 \\ 1/2 & 1/3 \end{bmatrix} = \begin{bmatrix} 3/4 & -1/3 \\ -1/2 & 2/3 \end{bmatrix}$$

de modo que la ecuación $(I - C)\mathbf{x} = \mathbf{d}$ se convierte en

$$\begin{bmatrix} 3/4 & -1/3 \\ -1/2 & 2/3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

En la práctica se reduciría por renglón la correspondiente matriz aumentada para determinar una solución. Sin embargo, en este caso es instructivo notar que la matriz de coeficientes $I - C$ es invertible y entonces aplicar el Teorema 3.7. Se calcula

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3/4 & -1/3 \\ -1/2 & 2/3 \end{bmatrix}^{-1} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 3/2 & 9/4 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

Dado que d_1, d_2 y todas las entradas de $(I - C)^{-1}$ son no negativas, también lo son x_1 y x_2. Por tanto, puede encontrarse una solución factible para *cualquier* vector de demanda distinto de cero.

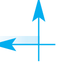
(b) En este caso,

$$I - C = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 2/3 \end{bmatrix} \quad \text{y} \quad (I - C)^{-1} = \begin{bmatrix} -4 & -6 \\ -6 & -6 \end{bmatrix}$$

de modo que

$$\mathbf{x} = (I - C)^{-1}\mathbf{d} = \begin{bmatrix} -4 & -6 \\ -6 & -6 \end{bmatrix}\mathbf{d}$$

Dado que todas las entradas de $(I - C)^{-1}$ son negativas, esto no producirá una solución factible para *cualquier* vector de demanda distinto de cero $\mathbf{d}$.



Motivada por el ejemplo 3.66, se tiene la siguiente definición. (Para dos matrices $m \times n$; $A = [a_{ij}]$ y $B = [b_{ij}]$, se escribirá $A \geq B$ si $a_{ij} \geq b_{ij}$ para todo i y j. De igual modo, puede definirse $A > B$, $A \leq B$, etcétera. Una matriz A se llama no negativa si $A \geq O$ y positiva si $A > O$.)

Definición Una matriz de consumo C se llama ***productiva*** si $I - C$ es invertible y $(I - C)^{-1} \geq O$.

Ahora se dan tres resultados que proporcionan criterios para que una matriz de consumo sea productiva.

Teorema 3.34 Sea C una matriz de consumo. Entonces C es productiva si y sólo si existe un vector de producción $\mathbf{x} \geq \mathbf{0}$ tal que $\mathbf{x} > C\mathbf{x}$.

Demostración Suponga que C es productiva. Entonces $I - C$ es invertible y $(I - C)^{-1} \geq O$. Sea

$$\mathbf{j} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

Entonces $\mathbf{x} = (I - C)^{-1}\mathbf{j} \geq 0$ y $(I - C)\mathbf{x} = \mathbf{j} > 0$. Por tanto, $\mathbf{x} - C\mathbf{x} > 0$ o, de manera equivalente, $\mathbf{x} > C\mathbf{x}$

Por el contrario, suponga que existe un vector $\mathbf{x} \geq 0$ tal que $\mathbf{x} > C\mathbf{x}$. Dado que $C \geq O$ y $C \neq O$, se tiene $\mathbf{x} > 0$ por el ejercicio 35. Más aún, debe existir un número real λ con $0 < \lambda < 1$ tal que $C\mathbf{x} < \lambda\mathbf{x}$. Pero entonces

$$C^2\mathbf{x} = C(C\mathbf{x}) \leq C(\lambda\mathbf{x}) = \lambda(C\mathbf{x}) < \lambda(\lambda\mathbf{x}) = \lambda^2\mathbf{x}$$



Por inducción, puede demostrarse que $0 \leq C^n\mathbf{x} < \lambda^n\mathbf{x}$ para toda $n \geq 0$. (Escriba los detalles de esta demostración por inducción.) Dado que $0 < \lambda < 1$, λ^n se aproxima a 0 cuando n se vuelve grande. Por tanto, cuando $n \to \infty$, $\lambda^n\mathbf{x} \to 0$ y en consecuencia $C^n\mathbf{x} \to 0$. Puesto que $\mathbf{x} > 0$, debe tener $C^n \to O$ cuando $n \to \infty$.

Ahora considere la ecuación matricial

$$(I - C)(I + C + C^2 + \ldots + C^{n-1}) = I - C^n$$

Cuando $n \to \infty$, $C^n \to O$, así que se tiene

$$(I - C)(I + C + C^2 + \dots) = I - O = I$$

Por tanto, $I - C$ es invertible, con su inverso dado por la serie matricial infinita $I + C + C^2 + \dots$. Dado que todos los términos en esta serie son no negativos, también se tiene

$$(I - C)^{-1} = I + C + C^2 + \dots \geq O$$

En consecuencia, C es productiva.

Comentarios

- La serie infinita $I + C + C^2 + \dots$ es el análogo matricial de la serie geométrica $1 + x + x^2 + \dots$. Es posible que el lector esté familiarizado con el hecho de que, para $|x| < 1$, $1 + x + x^2 + \dots = 1/(1 - x)$.
- Dado que el vector $C\mathbf{x}$ representa las cantidades consumidas por cada industria, la desigualdad $\mathbf{x} > C\mathbf{x}$ significa que existe cierto nivel de producción para el cual cada industria produce más de lo que consume.
- Para un planteamiento alternativo a la primera parte de la demostración del Teorema 3.34, vea el ejercicio 42 de la sección 4.6.

COROLARIO 3.35 Sea C una matriz de consumo. Si la suma de cada renglón de C es menor que 1, entonces C es productiva.

La palabra *corolario* proviene de la palabra latina *corollarium*, que se refiere a una guirnalda otorgada como recompensa. Por tanto, un corolario es una pequeña recompensa adicional que se tiene de un teorema.

Demostración Si

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$$

entonces $C\mathbf{x}$ es un vector que consiste de las sumas de renglón de C. Si cada suma de renglón de C es menor que 1, entonces se satisface la condición $\mathbf{x} > C\mathbf{x}$. Por tanto, C es productiva.

COROLARIO 3.36 Sea C una matriz de consumo. Si la suma de cada columna de C es menor que 1, entonces C es productiva.

Demostración Si cada suma de columna de C es menor que 1, entonces cada suma de columna de C^T es menor que 1. Por tanto, C^T es productiva, por el Corolario 3.35. En consecuencia, por los Teoremas 3.9(d) y 3.4,

$$((I - C)^{-1})^T = ((I - C)^T)^{-1} = (I^T - C^T)^{-1} = (I - C^T)^{-1} \geq O$$

Se tiene que $(I - C)^{-1} \geq O$ también y, por tanto, C es productiva.

En el ejercicio 52 de la sección 7.2 se le pedirá dar demostraciones alternativas a los Corolarios 3.35 y 3.36.

De la definición de matriz de consumo se tiene que la suma de la columna j es el valor monetario total de todas las entradas necesarias para producir un importe monetario de la output de la industria j; esto es, el ingreso de la industria j supera sus gastos. Se dice que tal industria es ***rentable***. En consecuencia, el Corolario 3.36 puede parafrasearse para enunciar que una matriz de consumo es productiva si todas las industrias son rentables.

P. H. Lewis, "On the Use of Matrices en Certain Population Mathematics", *Biometrika* **33** (1945), pp. 183-212.

Crecimiento poblacional

Uno de los modelos más populares de crecimiento poblacional es uno basado en matrices, introducido por primera vez por P. H. Leslie en 1945. El *modelo de Leslie* describe el crecimiento de la porción femenina de una población, la cual se supone tiene una vida máxima. Las hembras se dividen en clases etáreas (por edades), todas las cuales abarcan un número igual de años. Con datos acerca de las tasas de natalidad promedio y probabilidades de supervivencia de cada clase, el modelo es capaz de determinar el crecimiento de la población a lo largo del tiempo.

Ejemplo 3.67

Cierta especie de escarabajo alemán, llamado de Vollmar-Wasserman (o escarabajo VW, para abreviar), vive cuando mucho 3 años. Las hembras del escarabajo VW se dividen en tres clases etáreas de 1 año cada una: jóvenes (0-1 año), juveniles (1-2 años) y adultas (2-3 años). Las jóvenes no ponen huevos; cada juvenil produce un promedio de cuatro escarabajos hembras; y cada adulta produce un promedio de tres hembras.

La tasa de supervivencia para jóvenes es 50% (esto es: la probabilidad de que una joven sobreviva hasta volverse juvenil es 0.5), y la tasa de supervivencia de las juveniles es 25%. Suponga que comienza con una población de 100 escarabajos VW hembras: 40 jóvenes, 40 juveniles y 20 adultas. Prediga la población de escarabajos para cada uno de los siguientes 5 años.

Solución Después de 1 año, el número de jóvenes será el número producido durante dicho año:

$$40 \times 4 + 20 \times 3 = 220$$

El número de juveniles simplemente será el número de jóvenes que sobrevivió:

$$40 \times 0.5 = 20$$

Del mismo modo, el número de adultas será el número de juveniles que sobrevivió:

$$40 \times 0.25 = 10$$

Puede combinar esto en una sola ecuación matricial

$$\begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix}\begin{bmatrix} 40 \\ 40 \\ 20 \end{bmatrix} = \begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix}$$

o $L\mathbf{x}_0 = \mathbf{x}_1$, donde $\mathbf{x}_0 = \begin{bmatrix} 40 \\ 40 \\ 20 \end{bmatrix}$ es el vector de distribución de población inicial y $\mathbf{x}_1 = \begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix}$ es la distribución después de 1 año. Se ve que la estructura de la ecuación es exactamente igual que la de las cadenas de Markov: $\mathbf{x}_{k+1} = L\mathbf{x}_k$ para $k = 0, 1, 2, \ldots$ (aunque la interpretación es bastante diferente). Se tiene que es posible calcular por iteración vectores de distribución de población sucesivos. (También se tiene que $\mathbf{x}_k = L^k\mathbf{x}_0$ para $k = 0, 1, 2, \ldots$, como para las cadenas de Markov, pero aquí no se usará este hecho.)

Calcule

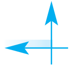
$$\mathbf{x}_2 = L\mathbf{x}_1 = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix}\begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 110 \\ 110 \\ 5 \end{bmatrix}$$

$$\mathbf{x}_3 = L\mathbf{x}_2 = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix}\begin{bmatrix} 110 \\ 110 \\ 5 \end{bmatrix} = \begin{bmatrix} 455 \\ 55 \\ 27.5 \end{bmatrix}$$

$$\mathbf{x}_4 = L\mathbf{x}_3 = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 455 \\ 55 \\ 27.5 \end{bmatrix} = \begin{bmatrix} 302.5 \\ 227.5 \\ 13.75 \end{bmatrix}$$

$$\mathbf{x}_5 = L\mathbf{x}_4 = \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix} \begin{bmatrix} 302.5 \\ 227.5 \\ 13.75 \end{bmatrix} = \begin{bmatrix} 951.2 \\ 151.2 \\ 56.88 \end{bmatrix}$$

Por tanto, el modelo predice que, después de 5 años, habrá aproximadamente 951 hembras jóvenes de escarabajo VW, 151 juveniles y 57 adultas. (*Nota:* podría argumentar que en cada paso habría que redondear al entero más cercano, por ejemplo, 28 adultos después del paso 3, lo que afectaría las iteraciones posteriores. Se eligió *no* hacer esto, pues, de cualquier forma, los cálculos sólo son aproximaciones y es mucho más fácil usar una calculadora o CAS si no redondea conforme avanza.)



La matriz L en el ejemplo 3.67 se llama ***matriz de Leslie***. En general, si se tiene una población con n clases etáreas de igual duración, L será una matriz de $n \times n$ con la siguiente estructura:

$$L = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{n-1} & b_n \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{n-1} & 0 \end{bmatrix}$$

Aquí, $b_1, b_2, \ldots$ son los *parámetros de natalidad* (b_i = número promedio de hembras producidas por cada hembra en la clase i) y $s_1, s_2, \ldots$ son las *probabilidades de supervivencia* (s_i = probabilidad de que una hembra en la clase i sobreviva en la clase $i + 1$).

¿Qué se hará con los cálculos? En general, la población de escarabajos parece aumentar, aunque existen algunas fluctuaciones, como una disminución de 250 a 225 del año 1 al año 2. La figura 3.22 muestra el cambio en la población en cada una de las tres clases etáreas y claramente muestra el crecimiento, con fluctuaciones.

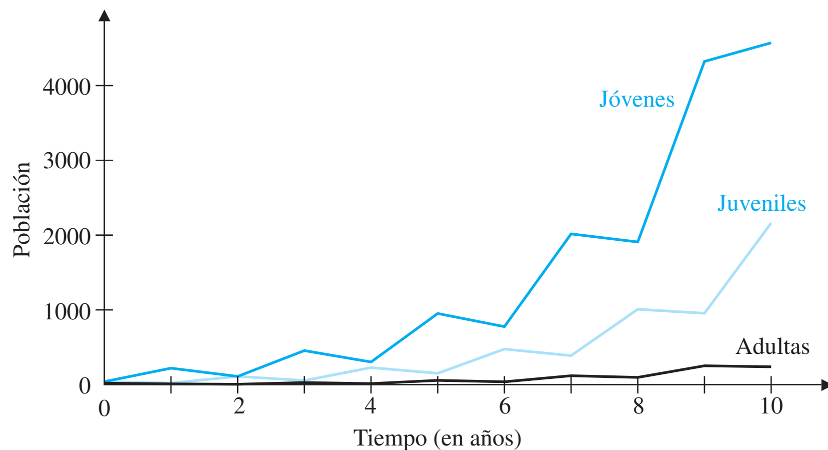


Figura 3.22

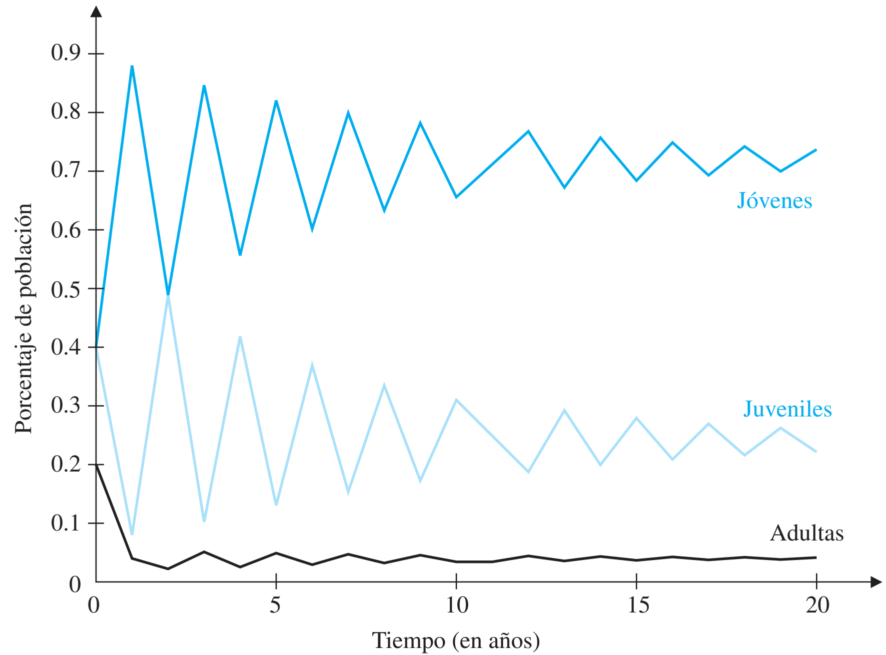


Figura 3.23

Si, en lugar de graficar la población *real* se grafica la población *relativa* en cada clase, surge un patrón diferente. Para hacer esto, es necesario calcular la fracción de la población en cada clase etárea en cada año; esto es, es necesario dividir cada vector de distribución por la suma de sus componentes. Por ejemplo, después de 1 año, se tiene

$$\frac{1}{250}\mathbf{x}_1 = \frac{1}{250}\begin{bmatrix} 220 \\ 20 \\ 10 \end{bmatrix} = \begin{bmatrix} 0.88 \\ 0.08 \\ 0.04 \end{bmatrix}$$

la cual dice que 88% de la población consiste de jóvenes, 8% de juveniles y 4% de adultas. Si este tipo de datos se grafican contra el tiempo, se obtiene una gráfica como la de la figura 3.23, que muestra claramente que la proporción de la población en cada clase se aproxima a un estado estable. Se evidencia que el vector de estado estacionario en este ejemplo es

$$\begin{bmatrix} 0.72 \\ 0.24 \\ 0.04 \end{bmatrix}$$

Esto es, a largo plazo, 72% de la población será de jóvenes, 24% de juveniles y 4% de adultas. (En otras palabras, la población se distribuye entre las tres clases etáreas en la razón 18 : 6 : 1.) En el capítulo 4 se verá cómo determinar esta razón con exactitud.

Grafos y digrafos

Existen muchas situaciones en las que es importante poder modelar las interrelaciones entre un conjunto finito de objetos. Por ejemplo, acaso quiera describir varios tipos de redes (carreteras que conectan ciudades, rutas de aerolíneas que conectan ciudades, ligas de comunicación que conectan satélites, etcétera) o relaciones entre grupos o individuos (relaciones de amistad en una sociedad, relaciones depredador-presa en un ecosistema,

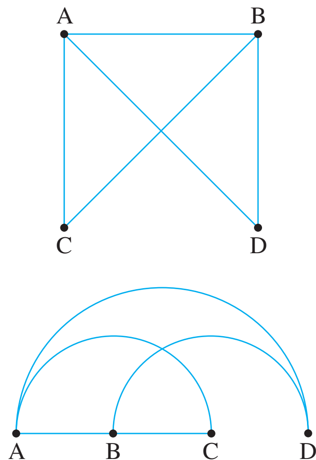


Figura 3.24
Dos representaciones del mismo grafo

relaciones de dominio en un deporte, etcétera). Los grafos son idealmente adecuados para modelar tales redes y relaciones, y es evidente que las matrices son una herramienta útil en su estudio.

Un ***grafo*** consiste de un conjunto finito de puntos (llamados *vértices*) y un conjunto finito de *aristas*, cada una de las cuales conecta dos vértices (no necesariamente distintos). Se dice que dos vértices son *adyacentes* si son los puntos finales de una arista. La figura 3.24 muestra un ejemplo del mismo grafo dibujado en dos formas diferentes. Los grafos son "iguales" en el sentido de que todo por lo que debe preocuparse es por las relaciones de adyacencia que identifican las aristas.

Es posible registrar la información esencial acerca de un grafo en una matriz y usar álgebra matricial para ayudar a responder ciertas preguntas acerca del grafo. Esto es particularmente útil si los grafos son grandes, pues las computadoras pueden manejar los cálculos muy rápidamente.

Definición Si G es un grafo con n vértices, entonces su ***matriz de adyacencia*** es la matriz A [o $A(G)$] de $n \times n$ definida por

$$a_{ij} = \begin{cases} 1 & \text{si existe una arista entre los vértices } i \text{ y } j \\ 0 & \text{de otro modo} \end{cases}$$

El término *vértice* viene del verbo latino *vertere*, que significa "volver". En el contexto de los grafos (y la geometría), un vértice es una esquina, un punto donde una arista "se vuelve" una arista diferente.

La figura 3.25 muestra un grafo y su matriz de adyacencia asociada.

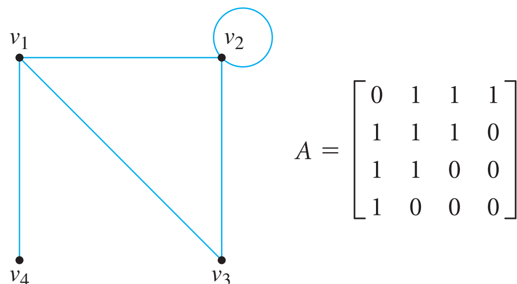


Figura 3.25
Un grafo con matriz de adyacencia A



Comentario Observe que la matriz de adyacencia de un grafo necesariamente es una matriz simétrica. (¿Por qué?) Note también que una entrada diagonal a_{ii} de A es cero a menos que haya un bucle en el vértice i. En algunas situaciones, un grafo puede tener más de una arista entre un par de vértices. En tales casos, puede tener sentido modificar la definición de la matriz de adyacencia de modo que a_{ij} sea igual al *número* de aristas entre los vértices i y j.

Una ***trayectoria*** en un grafo es una secuencia de aristas que permiten viajar de un vértice a otro de manera continua. La ***longitud*** de una trayectoria es el número de aristas que contiene, y a una trayectoria con k aristas se le denominará ***k*-*trayectoria***. Por ejemplo, en el grafo de la figura 3.25, $v_1v_3v_2v_1$ es una 3-trayectoria, y $v_4v_1v_2v_2v_1v_3$ es una 5-trayectoria. Note que la primera de éstas es *cerrada* (comienza y termina en el mismo vértice); a tal trayectoria se le llama ***circuito***. La segunda usa las aristas entre v_1 y v_2 dos veces; una trayectoria que *no* incluye la misma arista más de una vez se llama trayectoria ***simple***.

Es posible usar las potencias de la matriz de adyacencia de un grafo para obtener información acerca de las trayectorias de varias longitudes en el grafo. Considere el cuadrado de la matriz de adyacencia en la figura 3.25:

$$A^2 = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 2 & 3 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

¿Qué representan las entradas de A^2? Observe la entrada (2, 3). A partir de la definición de multiplicación matricial, se sabe que

$$(A^2)_{23} = a_{21}a_{13} + a_{22}a_{23} + a_{23}a_{33} + a_{24}a_{43}$$

La única manera en que esta expresión puede resultar en un número distinto de cero es si al menos uno de los productos $a_{2k}a_{k3}$ que constituyen la suma es distinto de cero. Pero $a_{2k}a_{k3}$ es distinto de cero si y sólo si tanto a_{2k} como a_{k3} son distintos de cero, lo que significa que hay una arista entre v_2 y v_k, así como una arista entre v_k y v_3. Por tanto, habrá una 2-trayectoria entre los vértices 2 y 3 (vía el vértice k). En el ejemplo, esto ocurre para $k =$ 1 y para $k = 2$, de modo que

$$\begin{aligned}(A^2)_{23} &= a_{21}a_{13} + a_{22}a_{23} + a_{23}a_{33} + a_{24}a_{43} \\ &= 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 0 + 0 \cdot 0 \\ &= 2\end{aligned}$$



la cual dice que hay dos 2-trayectorias entre los vértices 2 y 3. (Compruebe para ver que las entradas restantes de A^2 correctamente dan 2-trayectorias en el grafo.) El argumento que acaba de darse puede generalizarse para producir el siguiente resultado, cuya demostración se deja como ejercicio 72.

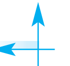
Si A es la matriz de adyacencia de un grafo G, entonces la entrada (i, j) de A^k es igual al número de k-trayectorias entre los vértices i y j.

Ejemplo 3.68 ¿Cuántas 3-trayectorias hay entre v_1 y v_2 en la figura 3.25?

Solución Se necesita la entrada (1, 2) de A^3, que es el producto punto del renglón 1 de A^2 y la columna 2 de A. El cálculo produce

$$(A^3)_{12} = 3 \cdot 1 + 2 \cdot 1 + 1 \cdot 1 + 0 \cdot 0 = 6$$

de modo que hay seis 3-trayectorias entre los vértices 1 y 2, lo que puede comprobarse con facilidad.



En muchas aplicaciones que pueden modelarse mediante un grafo, los vértices se ordenan mediante algún tipo de relación que impone una dirección a las aristas. Por ejemplo, las aristas dirigidas pueden usarse para representar rutas de una vía en un grafo que modela una red de transporte o relaciones depredador-presa en un grafo que modela un ecosistema. Un grafo con aristas dirigidas se llama ***digrafo***. La figura 3.26 muestra un ejemplo.

Una sencilla modificación de la definición de las matrices de adyacencia permite usarlas con digrafos.

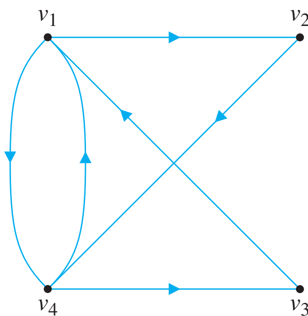


Figura 3.26
Un digrafo

Definición Si G es un digrafo con n vértices, entonces su ***matriz de adyacencia*** es la matriz A [o $A(G)$] de $n \times n$ definida por

$$a_{ij} = \begin{cases} 1 & \text{si existe una arista del vértice } i \text{ al vértice } j \\ 0 & \text{de otro modo} \end{cases}$$

Por tanto, la matriz de adyacencia para el digrafo en la figura 3.26 es

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$



No es de sorprender que la matriz de adyacencia de un digrafo no sea simétrica en general. (¿Cuándo lo sería?) El lector no debe tener dificultad para ver que A^k ahora contiene los números de k-trayectorias *dirigidas* entre vértices, donde se insiste que todas las aristas a lo largo de una trayectoria fluyen en la misma dirección. (Vea el ejercicio 72.) El siguiente ejemplo ofrece una aplicación de esta idea.

Ejemplo 3.69

Cinco tenistas (Djokovic, Federer, Nadal, Roddick y Safin) compiten en un torneo con el sistema *round-robin* en el que cada jugador compite con todos los demás una vez. El digrafo en la figura 3.27 resume los resultados. Una arista dirigida del vértice i al vértice j significa que el jugador i venció al jugador j. (Un digrafo en el que existe exactamente una arista dirigida entre cada par de vértices se llama ***torneo***.)

La matriz de adyacencia para el digrafo de la figura 3.27 es

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

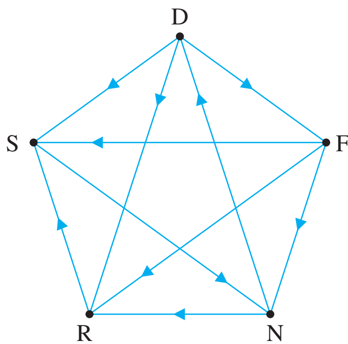


Figura 3.27
Un torneo

donde el orden de los vértices (y por tanto, los renglones y columnas de A) se determina alfabéticamente. Por tanto, Federer corresponde a renglón 2 y columna 2, por ejemplo.

Suponga que se quiere clasificar a los cinco jugadores con base en los resultados de sus partidos. Una forma de hacer esto puede ser contar el número de triunfos de cada jugador. Observe que el número de triunfos de cada jugador es justo la suma de las entradas en el renglón correspondiente; de manera equivalente, el vector que contiene todas las sumas de renglón está dado por el producto $A\mathbf{j}$, donde

$$\mathbf{j} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

En este caso se tiene

$$A\mathbf{j} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 2 \\ 1 \\ 1 \end{bmatrix}$$

que produce la siguiente clasificación:

Primero: Djokovic, Federer (empatados)
Segundo: Nadal
Tercero: Roddick, Safin (empatados)

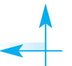
¿Los jugadores empatados en esta clasificación son igualmente fuertes? Djokovic puede argumentar que, puesto que venció a Federer, él merece el primer lugar. Roddick usaría el mismo tipo de argumento para romper el empate con Safin. Sin embargo, Safin podría argumentar que él tiene dos victorias "indirectas" porque venció a Nadal, quien venció a otros *dos*; más aún, él puede puntualizar que Roddick sólo tiene *una* victoria indirecta (sobre Safin, quien venció a Nadal).

Dado que en un grupo de empates puede no haber un jugador que venciera a todos los demás en el grupo, la notación de victorias indirectas parece más útil. Más aún, una victoria indirecta corresponde a una 2-trayectoria en el digrafo, de modo que puede usar el cuadrado de la matriz de adyacencia. Para calcular tanto los triunfos como las victorias indirectas para cada jugador, se necesitan las sumas de renglón de la matriz $A + A^2$, que están dadas por

$$(A + A^2)\mathbf{j} = \left(\begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 2 & 1 & 2 \\ 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 2 & 2 & 3 \\ 1 & 0 & 2 & 2 & 2 \\ 1 & 1 & 0 & 2 & 2 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 7 \\ 6 \\ 2 \\ 3 \end{bmatrix}$$

En consecuencia, los jugadores se clasificarían en el siguiente orden: Djokovic, Federer, Nadal, Safin, Roddick. Por desgracia, este planteamiento no garantiza romper todos los empates.



Códigos de corrección de error

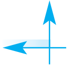
En la sección 1.4 se estudiaron ejemplos de códigos de detección de error. Ahora se aborda el problema de diseñar códigos que puedan *corregir* así como detectar ciertos tipos de errores. El mensaje será un vector $\mathbf{x}$ en $\mathbb{Z}_2^k$ para alguna k, y se le codificará usando

una transformación matricial $T : \mathbb{Z}_2^k \to \mathbb{Z}_2^n$ para algún $n > k$. El vector $T(\mathbf{x})$ se llamará ***vector código***. Un ejemplo simple servirá para ilustrar el enfoque a tomar, que es una generalización de los vectores de control de paridad del ejemplo 1.37.

Ejemplo 3.70

Suponga que el mensaje es un solo dígito binario: 0 o 1. Si el mensaje se codifica al simplemente repetirlo dos veces, entonces los vectores código son [0, 0] y [1, 1]. Este código puede detectar errores individuales. Por ejemplo, si se transmite [0, 0] y ocurre un error en el primer componente, entonces se recibe [1, 0] y se detecta un error, porque este no es un vector código legal. Sin embargo, el receptor no puede corregir el error, pues [1, 0] también sería resultado de un error en el segundo componente si se hubiese transmitido [1, 1].

Este problema se puede resolver al hacer los vectores código más largos: repetir el dígito de mensaje tres veces en lugar de dos. Por tanto, 0 y 1 se codifican como [0, 0, 0] y [1, 1, 1], respectivamente. Ahora, si ocurre un solo error, no sólo se puede detectar, sino también corregir. Por ejemplo, si se recibe [0, 1, 0], entonces se sabe que debe ser el resultado de un solo error en la transmisión de [0, 0, 0], pues un solo error en [1, 1, 1] no se habría producido.



Note que el código en el ejemplo 3.70 puede lograrse mediante una transformación matricial, aunque en una particularmente trivial. Sea $G = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ y defina $T : \mathbb{Z}_2 \to \mathbb{Z}_2^3$ mediante $T(\mathbf{x}) = G\mathbf{x}$. (Aquí los elementos de $\mathbb{Z}_2$ se consideran matrices de 1×1.) La matriz G se llama ***matriz generadora*** del código.

Para decir si un vector recibido es un vector código, se realiza no una, sino *dos* comprobaciones de paridad. Se requiere que el vector recibido $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$ satisfaga $c_1 = c_2 = c_3$. Estas ecuaciones pueden escribirse como un sistema lineal sobre $\mathbb{Z}_2$:

$$\begin{aligned} c_1 &= c_2 \\ c_1 &= c_3 \end{aligned} \quad \text{o} \quad \begin{aligned} c_1 + c_2 &= 0 \\ c_1 + c_3 &= 0 \end{aligned} \tag{1}$$

Si se hace $P = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, entonces (1) es equivalente a $P\mathbf{c} = \mathbf{0}$. La matriz P se llama ***matriz de control de paridad*** para el código. Observe que $PG = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = O$.

Para ver cómo estas matrices entran en juego para la corrección de errores, suponga que se envía 1 como $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = [1 \quad 1 \quad 1]^T$, pero un solo error hace que se reciba como

$\mathbf{c}' = [1 \quad 0 \quad 1]^T$. Se calcula

$$P\mathbf{c}' = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \mathbf{0}$$

de modo que se sabe que $\mathbf{c}'$ no puede ser un vector código. ¿Dónde está el error? Note que $P\mathbf{c}'$ es la segunda columna de la matriz de control de paridad P; ésta dice que el error está en el segundo componente de $\mathbf{c}'$ (lo que se probará adelante en el Teorema 3.37) y permite corregir el error. (Desde luego, en este ejemplo podría encontrar el error más rápidamente sin usar matrices, pero la idea es útil.)

Para generalizar las ideas en el último ejemplo, se hacen las siguientes definiciones.

Definiciones Si $k < n$, entonces cualquier matriz de $n \times k$ de la forma $G = \begin{bmatrix} I_k \\ A \end{bmatrix}$, donde A es una matriz de $(n - k) \times k$ sobre $\mathbb{Z}_2$, se llama ***matriz generadora estándar*** para un ***código binario*** (n, k) $T : \mathbb{Z}_2^k \to \mathbb{Z}_2^n$. Cualquier matriz de $(n - k) \times n$ de la forma $P = [B \; I_{n-k}]$, donde B es una matriz de $(n - k) \times k$ sobre $\mathbb{Z}_2$, se llama ***matriz de comprobación de paridad estándar***. Se dice que el código tiene ***longitud*** n y ***dimensión*** k.

He aquí lo que se necesita saber: (a) ¿Cuándo es G la matriz generadora estándar para un código binario *de corrección de error*? (b) Dada G, ¿cómo se encuentra una matriz de control de paridad estándar asociada P? Es evidente que las respuestas son bastante sencillas, como se muestra mediante el siguiente teorema.

Teorema 3.37

Si $G = \begin{bmatrix} I_k \\ A \end{bmatrix}$ es una matriz generadora estándar y $P = [B \quad I_{n-k}]$ es una matriz de control de paridad estándar, entonces P es la matriz de control de paridad asociada con G, si y sólo si $A = B$. El correspondiente código binario (n, k) es de corrección de error (sencillo) si y sólo si las columnas de P son distintas de cero y diferentes.

Antes de demostrar el teorema, considere otro ejemplo menos trivial que lo ilustra.

Ejemplo 3.71

Suponga que quiere diseñar un código de corrección de error que use tres ecuaciones de control de paridad. Dado que dichas ecuaciones dan lugar a los renglones de P, se tiene $n - k = 3$ y $k = n - 3$. Los vectores de mensaje vienen de $\mathbb{Z}_2^k$, así que se buscaría que k (y por tanto n) fuese tan largo como sea posible, con la finalidad de poder transmitir tanta información como sea posible. Por el Teorema 3.37, las n columnas de P deben ser distintas de cero y diferentes así que el máximo ocurre cuando ellas consisten de todos los $2^3 - 1 = 7$ vectores distintos de cero de $\mathbb{Z}_2^{n-k} = \mathbb{Z}_2^3$. Uno de tales candidatos es

$$P = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Esto significa que

$$A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

y por tanto, por el Teorema 3.37, una matriz generadora estándar para este código es

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

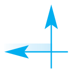
Como ejemplo de cómo funciona la matriz generadora, suponga que se codifica $\mathbf{x} = [0 \quad 1 \quad 0 \quad 1]^T$ para obtener el vector código

$$\mathbf{c} = G\mathbf{x} = [0 \quad 1 \quad 0 \quad 1 \quad 0 \quad 1 \quad 0]^T$$

Si se recibe este vector, se ve que es correcto, pues $P\mathbf{c} = \mathbf{0}$. Por otra parte, si se recibe $\mathbf{c}' = [0 \quad 1 \quad 1 \quad 1 \quad 0 \quad 1 \quad 0]^T$ se calcula

$$P\mathbf{c}' = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

que se reconoce como la columna 3 de P. Por tanto, el error está en el tercer componente de $\mathbf{c}'$, y al cambiarlo se recupera el vector código correcto $\mathbf{c}$. También se sabe que los primeros cuatro componentes de un vector código son el vector mensaje original, así que en este caso se decodifica $\mathbf{c}$ para obtener el original $\mathbf{x} = [0 \quad 1 \quad 0 \quad 1]^T$.



El código del ejemplo 3.71 se llama código de Hamming $(7, 4)$. Cualquier código binario construido de esta forma se llama ***código de Hamming* (n, k)**. Observe que, por construcción, un código de Hamming (n, k) tiene $n = 2^{n-k} - 1$.

Demostración del Teorema 3.37 (A lo largo de esta demostración se denota mediante $\mathbf{a}_i$ la i-ésima columna de una matriz A.) Con P y G como en el enunciado del teorema, primero suponga que son matrices de control de paridad estándar y generadora para el mismo código binario. Por tanto, para toda $\mathbf{x}$ en $\mathbb{Z}_2^k$, $PG\mathbf{x} = \mathbf{0}$. En términos de multiplicación de bloques,

$$[B \quad I]\begin{bmatrix} I \\ A \end{bmatrix}\mathbf{x} = \mathbf{0} \quad \text{para toda } \mathbf{x} \text{ en } \mathbb{Z}_2^k$$

De manera equivalente, para toda $\mathbf{x}$ en $\mathbb{Z}_2^k$, se tiene

$$B\mathbf{x} + A\mathbf{x} = (B + A)\mathbf{x} = (BI + IA)\mathbf{x} = [B \quad I]\begin{bmatrix} I \\ A \end{bmatrix}\mathbf{x} = \mathbf{0}$$

o

$$B\mathbf{x} = A\mathbf{x}$$

Si ahora toma $\mathbf{x} = \mathbf{e}_i$, el i-ésimo vector base estándar en $\mathbb{Z}_2^k$, se ve que

$$\mathbf{b}_i = B\mathbf{e}_i = A\mathbf{e}_i = \mathbf{a}_i \quad \text{para toda } i$$

Por tanto, $B = A$.

Por el contrario, es fácil comprobar que si $B = A$, entonces $PG\mathbf{x} = \mathbf{0}$ para toda $\mathbf{x}$ en $\mathbb{Z}_2^k$ (vea el ejercicio 92).

Para ver que tal par determina un código de corrección de error si las columnas de P son distintas de cero y diferentes, sea $\mathbf{x}$ un vector mensaje en $\mathbb{Z}_2^k$ y sea el correspondiente vector código $\mathbf{c} = G\mathbf{x}$. Entonces $P\mathbf{c} = \mathbf{0}$. Suponga que hay un error en el i-ésimo componente, que resulta en el vector $\mathbf{c}'$. Se tiene que $\mathbf{c}' = \mathbf{c} + \mathbf{e}_i$. Ahora calcule

$$P\mathbf{c}' = P(\mathbf{c} + \mathbf{e}_i) = P\mathbf{c} + P\mathbf{e}_i = \mathbf{0} + \mathbf{p}_i = \mathbf{p}_i$$

que localiza el error en el i-ésimo componente.

Por otra parte, si $\mathbf{p}_i = \mathbf{0}$, entonces un error en el i-ésimo componente no se detectará (es decir, $P\mathbf{c}' = \mathbf{0}$), y si $\mathbf{p}_i = \mathbf{p}_j$, entonces no se puede determinar si ocurrió un error en el i-ésimo o el j-ésimo componente (ejercicio 93).

Richard W. Hamming (1915–1998) recibió su doctorado en matemáticas de la Universidad de Illinois en Urbana-Champaign en 1942. Sus intereses de investigación en matemáticas fueron los campos de las ecuaciones diferenciales y el análisis numérico. De 1946 a 1976, trabajó en los Laboratorios Bell, luego de lo cual se unió al cuerpo docente en la Escuela de Posgraduados Navales de Estados Unidos, en Monterey, California. En 1950 publicó su ensayo fundamental acerca de los códigos de corrección de errores, lo que proporcionó una construcción explícita para los códigos óptimos que Claude Shannon probó teóricamente posibles en 1948.

A continuación se resumen las ideas principales de esta sección.

1. Para $n > k$, una matriz G de $n \times k$ y una matriz P de $(n - k) \times n$ (con entradas en $\mathbb{Z}_2$) son una matriz generadora estándar y una matriz de control de paridad estándar, respectivamente, para un código binario (n, k) si y sólo si, en forma de bloque,

$$G = \begin{bmatrix} I_k \\ A \end{bmatrix} \quad \text{y} \quad P = [A \quad I_{n-k}]$$

para alguna matriz A de $(n - \text{k}) \times k$ sobre $\mathbb{Z}_2$.
2. G codifica un vector mensaje $\mathbf{x}$ en $\mathbb{Z}_2^k$ como un vector código $\mathbf{c}$ en $\mathbb{Z}_2^n$ vía $\mathbf{c} = G\mathbf{x}$.
3. G es de corrección de error si y sólo si las columnas de P son distintas de cero y diferentes. Un vector $\mathbf{c}'$ en $\mathbb{Z}_2^n$ es un vector código si y sólo si $P\mathbf{c}' = \mathbf{0}$. En este caso, el correspondiente vector mensaje es el vector $\mathbf{x}$ en $\mathbb{Z}_2^k$ que consiste de los primeros k componentes de $\mathbf{c}'$. Si $P\mathbf{c}' \neq \mathbf{0}$, entonces $\mathbf{c}'$ no es un vector código y $P\mathbf{c}'$ es una de las columnas de P. Si $P\mathbf{c}'$ es la i-ésima columna de P, entonces el error está en el i-ésimo componente de $\mathbf{c}'$ y se puede recuperar el vector código correcto (y en consecuencia el mensaje) al cambiar este componente.

Ejercicios 3.7

Cadenas de Markov

En los ejercicios 1-4, sea $P = \begin{bmatrix} 0.5 & 0.3 \\ 0.5 & 0.7 \end{bmatrix}$ *la matriz de transición para una cadena de Markov con dos estados. Sea* $\mathbf{x}_0 = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$ *el vector de estado inicial para la población.*

1. Calcule $\mathbf{x}_1$ y $\mathbf{x}_2$.
2. ¿Qué proporción de la población del estado 1 estará en el estado 2 después de dos pasos?
3. ¿Qué proporción de la población del estado 2 estará en el estado 2 después de dos pasos?
4. Encuentre el vector de estado estacionario.

En los ejercicios 5-8, sea $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{3} & 0 \end{bmatrix}$ *la matriz de transición para una cadena de Markov con tres estados. Sea* $\mathbf{x}_0 = \begin{bmatrix} 120 \\ 180 \\ 90 \end{bmatrix}$ *el vector de estado inicial para la población.*

5. Calcule $\mathbf{x}_1$ y $\mathbf{x}_2$.
6. ¿Qué proporción de la población del estado 1 estará en el estado 1 después de dos pasos?
7. ¿Qué proporción de la población del estado 2 estará en el estado 3 después de dos pasos?
8. Encuentre el vector de estado estacionario.
9. Suponga que el clima en una región particular se comporta de acuerdo con una cadena de Markov. Específicamente, suponga que la probabilidad de que mañana será un día húmedo es de 0.662 si hoy es húmedo y de 0.250 si hoy es seco. La probabilidad de que mañana sea un día seco es de 0.750 si hoy es seco y de 0.338 si hoy es húmedo. [Este ejercicio se basa en un estudio real de lluvias en Tel Aviv durante un periodo de 27 años. Vea K. R. Gabriel y J. Neumann, "A Markov Chain Model for Daily Rainfall Occurrence at Tel Aviv", *Quarterly Journal of the Royal Meteorological Society, 88* (1962), pp. 90-95.]
   (a) Escriba la matriz de transición para esta cadena de Markov.
   (b) Si el lunes es un día seco, ¿cuál es la probabilidad de que el miércoles sea húmedo?
   (c) A largo plazo, ¿cuál será la distribución de días húmedos y secos?
10. Se han acumulado datos acerca de las estaturas de niños en relación con sus padres. Suponga que las probabilidades de que un padre alto tenga un hijo alto, de mediana estatura o bajo son 0.6, 0.2 y 0.2, respectivamente; las probabilidades de que un padre de talla media tenga un hijo alto, de estatura media o bajo son 0.1, 0.7 y 0.2, respectivamente; y las probabilidades de que un padre bajo tenga un hijo alto, mediano o bajo son 0.2, 0.4 y 0.4, respectivamente.
    (a) Escriba la matriz de transición para esta cadena de Markov.
    (b) ¿Cuál es la probabilidad de que una persona baja tenga un nieto alto?
    (c) Si 20% de la población actual es alta, 50% es de estatura media y 30% es baja, ¿cuál será la distribución en tres generaciones?
    (d) Si los datos del inciso (c) no cambian con el tiempo, ¿qué proporción de la población será alta, de estatura mediana y baja a largo plazo?
11. Un estudio de cultivos de pino piñonero en el sureste estadounidense, de 1940 a 1947, planteó la hipótesis que la producción de pinos seguía una cadena de Markov. [Vea D. H. Thomas, "A Computer Simulation Model of Great Basin Shoshonean Subsistence and Settlement Patterns", en D. L. Clarke ed., *Models in Archaeology* (Londres: Methuen, 1972).] Los datos sugirieron que si la cosecha de un año era buena, entonces la probabilidades de que el cultivo del siguiente año fuera bueno, adecuado o pobre eran de 0.08, 0.07 y 0.85, respectivamente; si la cosecha de un año era adecuada, entonces las probabilidades de que la cosecha del siguiente año fuera buena, adecuada y pobre eran de 0.09, 0.11 y 0.80, respectivamente; si la cosecha de un año era pobre, entonces las probabilidades de que la cosecha del siguiente año fuera buena, adecuada o pobre eran de 0.11, 0.05 y 0.84, respectivamente.
    (a) Escriba la matriz de transición para esta cadena de Markov.
    (b) Si la cosecha de pino piñonero fue buena en 1940, encuentre las probabilidades de una buena cosecha en los años 1941 a 1945.
    (c) A largo plazo, ¿qué proporción de las cosechas será buena, adecuada y pobre?
12. Se han programado robots para recorrer el laberinto que se muestra en la figura 3.28 y en cada unión eligen cuál camino seguir en forma aleatoria.

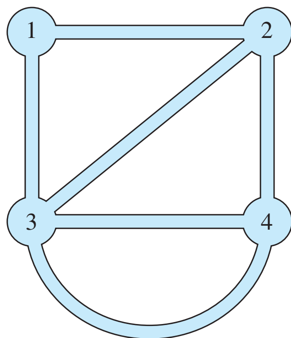


Figura 3.28

(a) Construya la matriz de transición para la cadena de Markov que modela esta situación.
(b) Suponga que comienza con 15 robots en cada unión. Encuentre la distribución de estado estacionario de robots. (Suponga que a cada robot le toma la misma cantidad de tiempo recorrer la distancia entre dos uniones adyacentes.)

13. Sea $\mathbf{j}$ un vector renglón que consiste por completo de números 1. Demuestre que una matriz P no negativa es una matriz estocástica si y sólo si $\mathbf{j}P = \mathbf{j}$.

14. (a) Demuestre que el producto de dos matrices estocásticas de 2×2 también es una matriz estocástica.
(b) Demuesrtre que el producto de dos matrices estocásticas de $n \times n$ también es una matriz estocástica.
(c) Si una matriz estocástica P de 2×2 es invertible, demuestre que P^{-1} también es una matriz estocástica.

Suponga que se quiere conocer el número promedio (o esperado) de pasos que se darán para ir del estado i al estado j en una cadena de Markov. Se puede demostrar que el siguiente cálculo responde esta pregunta: borre el renglón j-ésimo y la columna j-ésima de la matriz de transición P para obtener una nueva matriz Q. (Mantenga los renglones y columnas de Q etiquetadas como lo estaban en P.) El número esperado de pasos del estado i al estado j está dado por la suma de las entradas en la columna de $(I - Q)^{-1}$ *marcada i.*

15. En el ejercicio 9, si el lunes es un día seco, ¿cuál es el número esperado de días hasta un día húmedo?

16. En el ejercicio 10, ¿cuál es el número esperado de generaciones hasta que una persona baja tenga un descendiente alto?

17. En el ejercicio 11, si la cosecha de pino piñonero es adecuada un año, ¿cuál es el número esperado de años que deben pasar hasta que ocurre una buena cosecha?

18. En el ejercicio 12, a partir de cada una de las otras uniones, ¿cuál es el número esperado de movimientos hasta que un robot llegue a la unión 4?

Modelos económicos lineales

En los ejercicios 19-26, determine cuáles de las matrices son de intercambio. Para las que lo sean, encuentre un vector de precio no negativo que satisfaga la ecuación (1).

19. $\begin{bmatrix} 1/2 & 1/4 \\ 1/2 & 3/4 \end{bmatrix}$

20. $\begin{bmatrix} 1/3 & 2/3 \\ 1/2 & 1/2 \end{bmatrix}$

21. $\begin{bmatrix} 0.4 & 0.7 \\ 0.6 & 0.4 \end{bmatrix}$

22. $\begin{bmatrix} 0.1 & 0.6 \\ 0.9 & 0.4 \end{bmatrix}$

23. $\begin{bmatrix} 1/3 & 0 & 0 \\ 1/3 & 3/2 & 0 \\ 1/3 & -1/2 & 1 \end{bmatrix}$

24. $\begin{bmatrix} 1/2 & 1 & 0 \\ 0 & 0 & 1/3 \\ 1/2 & 0 & 2/3 \end{bmatrix}$

25. $\begin{bmatrix} 0.3 & 0 & 0.2 \\ 0.3 & 0.5 & 0.3 \\ 0.4 & 0.5 & 0.5 \end{bmatrix}$

26. $\begin{bmatrix} 0.50 & 0.70 & 0.35 \\ 0.25 & 0.30 & 0.25 \\ 0.25 & 0 & 0.40 \end{bmatrix}$

En los ejercicios 27-30, determine si la matriz de consumo dada es productiva.

27. $\begin{bmatrix} 0.2 & 0.3 \\ 0.5 & 0.6 \end{bmatrix}$

28. $\begin{bmatrix} 0.20 & 0.10 & 0.10 \\ 0.30 & 0.15 & 0.45 \\ 0.15 & 0.30 & 0.50 \end{bmatrix}$

29. $\begin{bmatrix} 0.35 & 0.25 & 0 \\ 0.15 & 0.55 & 0.35 \\ 0.45 & 0.30 & 0.60 \end{bmatrix}$

30. $\begin{bmatrix} 0.2 & 0.4 & 0.1 & 0.4 \\ 0.3 & 0.2 & 0.2 & 0.1 \\ 0 & 0.4 & 0.5 & 0.3 \\ 0.5 & 0 & 0.2 & 0.2 \end{bmatrix}$

En los ejercicios 31-34, se proporcionan una matriz de consumo C y un vector de demanda **d**. *En cada caso, encuentre un vector de producción factible* **x** *que satisfaga la ecuación (2).*

31. $C = \begin{bmatrix} 1/2 & 1/4 \\ 1/2 & 1/2 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$

32. $C = \begin{bmatrix} 0.1 & 0.4 \\ 0.3 & 0.2 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

33. $C = \begin{bmatrix} 0.5 & 0.2 & 0.1 \\ 0 & 0.4 & 0.2 \\ 0 & 0 & 0.5 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$

CAS **34.** $C = \begin{bmatrix} 0.1 & 0.4 & 0.1 \\ 0 & 0.2 & 0.2 \\ 0.3 & 0.2 & 0.3 \end{bmatrix}, \mathbf{d} = \begin{bmatrix} 1.1 \\ 3.5 \\ 2.0 \end{bmatrix}$

35. Sea A una matriz de $n \times n$ y $A \geq O$. Suponga que $A\mathbf{x} < \mathbf{x}$ para alguna $\mathbf{x}$ en $\mathbb{R}^n$, $\mathbf{x} \geq \mathbf{0}$. Demuestre que $\mathbf{x} > \mathbf{0}$.

36. Sean A, B, C y D matrices de $n \times n$ y $\mathbf{x}, \mathbf{y}$ vectores en $\mathbb{R}^n$. demuestre las siguientes desigualdades:

(a) Si $A \geq B \geq O$ y $C \geq D \geq O$, entonces $AC \geq BD \geq O$.

(b) Si $A > B$ y $\mathbf{x} \geq \mathbf{0}$, $\mathbf{x} \neq \mathbf{0}$, entonces $A\mathbf{x} > B\mathbf{x}$.

Crecimiento poblacional

37. Una población con tres clases etáreas tiene una matriz de Leslie $L = \begin{bmatrix} 1 & 1 & 3 \\ 0.7 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$. Si el vector de población inicial es $\mathbf{x}_0 = \begin{bmatrix} 100 \\ 100 \\ 100 \end{bmatrix}$, calcule $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$.

38. Una población con cuatro clases etáreas tiene una matriz de Leslie $L = \begin{bmatrix} 0 & 1 & 2 & 5 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 \\ 0 & 0 & 0.3 & 0 \end{bmatrix}$. Si el vector de población inicial es $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 10 \\ 10 \\ 10 \end{bmatrix}$, calcule $\mathbf{x}_1$, $\mathbf{x}_2$ y $\mathbf{x}_3$.

39. Cierta especie con dos clases etáreas de 1 año de duración tiene una probabilidad de supervivencia de 80% de la clase 1 a la clase 2. Evidencia empírica demuestra que, en promedio, cada hembra pare cinco hembras por año. En consecuencia, dos posibles matrices de Leslie son

$$L_1 = \begin{bmatrix} 0 & 5 \\ 0.8 & 0 \end{bmatrix} \quad \text{y} \quad L_2 = \begin{bmatrix} 4 & 1 \\ 0.8 & 0 \end{bmatrix}$$

(a) A partir de $\mathbf{x}_0 = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$, calcule $\mathbf{x}_1, \ldots, \mathbf{x}_{10}$ en cada caso.

(b) Para cada caso, grafique el tamaño relativo de cada clase etárea contra el tiempo (como en la figura 3.23). ¿Qué sugieren sus gráficas?

40. Suponga que la matriz de Leslie para el escarabajo VW es $L = \begin{bmatrix} 0 & 0 & 20 \\ 0.1 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$. A partir de un $\mathbf{x}_0$ arbitrario, determine el comportamiento de esta población.

41. Suponga que la matriz de Leslie para el escarabajo VW es $L = \begin{bmatrix} 0 & 0 & 20 \\ s & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$. Investigue el efecto de variar la probabilidad de supervivencia s de los escarabajos jóvenes.

CAS **42.** Los caribúes se encuentran principalmente en las provincias occidentales de Canadá y el noroeste estadounidense. La vida promedio de una hembra es de alrededor de 14 años. Las tasas de natalidad y supervivencia para cada grupo etáreo se proporcionan en la tabla 3.4, que muestra que las hembras de caribú no dan a luz durante sus primeros 2 años y paren aproximadamente una cría por año durante sus años intermedios. La tasa de mortalidad para las crías jóvenes es muy alta.

Tabla 3.4

Edad (años)	Tasa de natalidad	Tasa de supervivencia
0–2	0.0	0.3
2–4	0.4	0.7
4–6	1.8	0.9
6–8	1.8	0.9
8–10	1.8	0.9
10–12	1.6	0.6
12–14	0.6	0.0

Los números de caribúes reportados en el Jasper National Park de Alberta en 1990 se muestran en la tabla 3.5. Con un CAS, prediga la población de caribúes para 1992 y 1994. Luego proyecte la población para los años 2010 y 2020. ¿Qué concluye? (¿Qué su-

Tabla 3.5 Población de caribués en el Jasper National Park, 1990

Edad (años)	Número
0–2	10
2–4	2
4–6	8
6–8	5
8–10	12
10–12	0
12–14	1

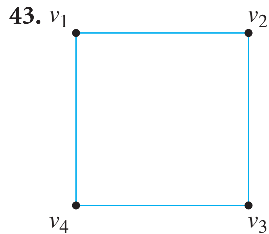
Fuente: World Wildlife Fund Canada

posiciones hace este modelo y cómo podría mejorarlo?)

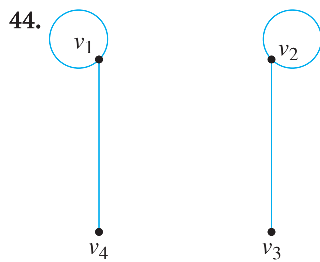
Grafos y digrafos

En los ejercicios 43-46, determine la matriz de adyacencia del grafo dado.

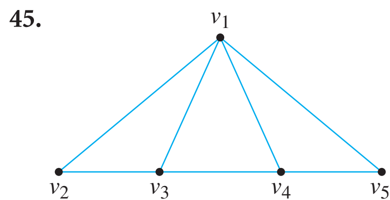
43.



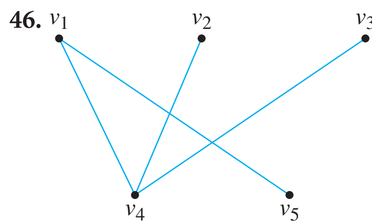
44.



45.



46.



En los ejercicios 47-50, dibuje un grafo que tenga la matriz de adyacencia dada.

47. $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$

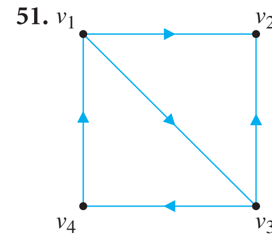
48. $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

49. $\begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \end{bmatrix}$

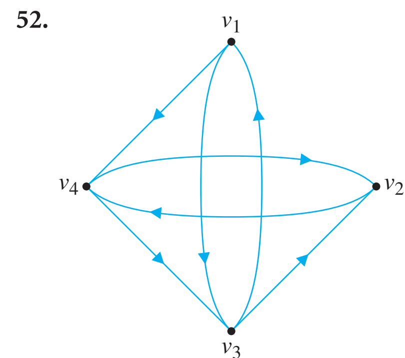
50. $\begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}$

En los ejercicios 51-54, determine la matriz de adyacencia del digrafo dado.

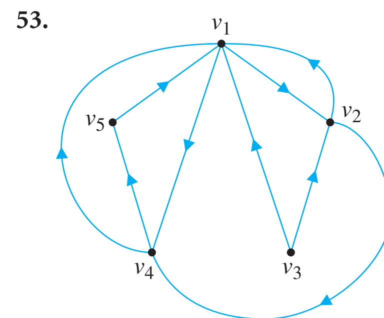
51.



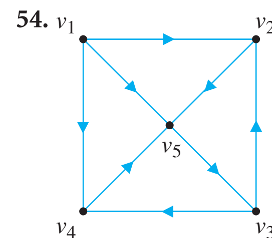
52.



53.



54.



En los ejercicios 55-58, dibuje un digrafo que tenga la matriz de adyacencia dada.

55. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}$ **56.** $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

57. $\begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}$ **58.** $\begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 59-66, use potencias de matrices de adyacencia para determinar el número de trayectorias de la longitud especificada entre los vértices dados.

59. Ejercicio 48, longitud 2, v_1 y v_2

60. Ejercicio 50, longitud 2, v_1 y v_2

61. Ejercicio 48, longitud 3, v_1 y v_3

62. Ejercicio 50, longitud 4, v_1 y v_2

63. Ejercicio 55, longitud 2, v_1 a v_3

64. Ejercicio 55, longitud 3, v_4 a v_1

65. Ejercicio 58, longitud 3, v_4 a v_1

66. Ejercicio 58, longitud 4, v_1 a v_4

67. Sea A la matriz de adyacencia de un grafo G.
- **(a)** Si el renglón i de A es todo ceros, ¿qué implica esto acerca de G?
- **(b)** Si la columna j de A es toda ceros, ¿qué implica esto acerca de G?

68. Sea A la matriz de adyacencia de un digrafo D.
- **(a)** Si el renglón i de A^2 es todo ceros, ¿qué implica esto acerca de D?
- **(b)** Si la columna j de A^2 es toda ceros, ¿qué implica esto acerca de D?

69. La figura 3.29 es el digrafo de un torneo con seis jugadores, P_1 a P_6. Con matrices de adyacencia, clasifique a los jugadores primero al determinar sólo victorias, y luego al usar la noción de victorias combinadas y victorias indirectas, como en el ejemplo 3.69.

70. La figura 3.30 es un digrafo que representa una cadena alimenticia en un pequeño ecosistema. Una arista dirigida de a a b indica que a tiene a b como fuente de alimento. Construya la matriz de adyacencia A para este digrafo y úsela para responder las siguientes preguntas.
- **(a)** ¿Cuál especie tiene las fuentes de alimento más directas? ¿Cómo muestra A esto?

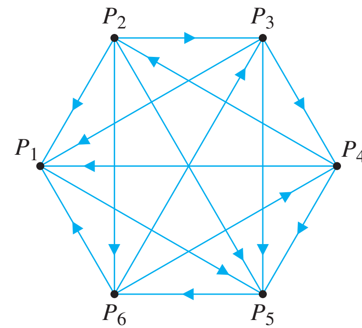


Figura 3.29

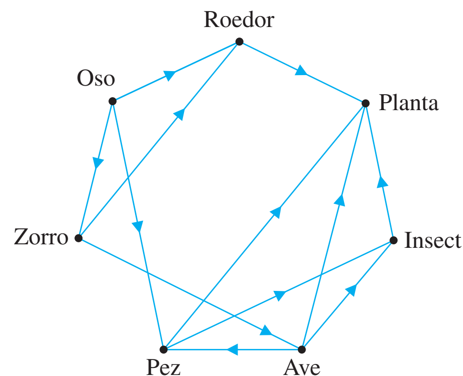


Figura 3.30

- **(b)** ¿Cuál especie es una fuente directa de alimento para la mayoría de las otras especies? ¿Cómo muestra A esto?
- **(c)** Si a come a b y b come a c, se dice que a tiene a c como una fuente indirecta de alimento. ¿Cómo puede usar A para determinar cuál especie tiene más fuentes indirectas de alimento? ¿Cuál especie tiene más fuentes de alimento directas e indirectas combinadas?
- **(d)** Suponga que los contaminantes matan a las plantas en esta cadena alimenticia y quiere determinar el efecto que tendrá este cambio sobre el ecosistema. Construya una nueva matriz de adyacencia A^* a partir de A, al borrar el renglón y la columna correspondientes a planta. Repita los incisos (a) a (c) y determine cuáles especies son las más y menos afectadas por el cambio.
- **(e)** ¿Cuál será el efecto a largo plazo de la contaminación? ¿Qué cálculos matriciales demostrarán esto?

71. Cinco personas se conectan todas mediante correo electrónico. Siempre que una de ellas escucha un trozo de chismorreo interesante, lo envía por correo electrónico a alguien más en el grupo, de acuerdo con la tabla 3.6.
- **(a)** Dibuje el digrafo que modela esta “red de chismorreo” y encuentre su matriz de adyacencia A.

Tabla 3.6

Remitente	Destinatarios
Ann	Carla, Ehaz
Bert	Carla, Dana
Carla	Ehaz
Dana	Ann, Carla
Ehaz	Bert

(**b**) Defina un *paso* como el tiempo que una persona tarda en enviar correo electrónico a todos los de su lista. (Por tanto, en un paso, el chisme va de Ann tanto a Carla como a Ehaz.) Si Bert escucha un rumor, ¿cuántos pasos transcurrirán para que todos los demás escuchen el rumor? ¿Qué cálculo matricial revela esto?

(**c**) Si Ann escucha un rumor, ¿cuántos pasos transcurrirán para que los demás escuchen el rumor? ¿Qué cálculo matricial revela esto?

(**d**) En general, si A es la matriz de adyacencia de un digrafo, ¿cómo puede decir si el vértice i está conectado al vértice j mediante una trayectoria (de cierta longitud)?

[La red de chismorreo en este ejercicio recuerda la noción de “seis grados de separación” (que se encuentra en la obra y la película del mismo nombre), que sugiere que cualesquiera dos personas están conectadas mediante una trayectoria de conocidos cuya longitud es cuando mucho 6. El juego “Six Degrees of Kevin Bacon” asevera más frívolamente que todos los actores están conectados al actor Kevin Bacon en tal forma.]

72. Sea A la matriz de adyacencia de un grafo G.

(**a**) Por inducción, pruebe que, para toda $n \geq 1$, la entrada (i, j) de A^n es igual al número de n-trayectorias entre los vértices i y j.

(**b**) ¿Cómo deben modificarse el enunciado y la demostración del inciso (a) si G es un digrafo?

73. Si A es la matriz de adyacencia de un digrafo G, ¿qué representa la entrada (i, j) de AA^T si $i \neq j$?

*Un grafo se llama **bipartita** si sus vértices pueden subdividirse en dos conjuntos U y V tales que toda arista tiene un punto final en U y el otro punto final en V. Por ejemplo, el grafo del ejercicio 46 es bipartita con $U = \{v_1, v_2, v_3\}$ y $V = \{v_4, v_5\}$. En los ejercicios 74-77, determine si un grafo con la matriz de adyacencia dada es bipartita.*

74. La matriz de adyacencia en el ejercicio 47

75. La matriz de adyacencia en el ejercicio 50

76. La matriz de adyacencia en el ejercicio 49

77. $\begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}$

78. (**a**) Demuestre que un grafo es bipartita si y sólo si sus vértices pueden etiquetarse de modo que su matriz de adyacencia pueda particionarse como

$$A = \left[\begin{array}{c|c} O & B \\ \hline B^T & O \end{array}\right]$$

(**b**) Con el resultado del inciso (a), demuestre que un grafo bipartita no tiene circuitos de longitud impar.

Códigos de corrección de error

79. Suponga que los cuatro vectores en $\mathbb{Z}_2^2$ se codifican al repetir el vector dos veces. Por tanto, se tiene

$$[0, 0] \rightarrow [0, 0, 0, 0]$$
$$[0, 1] \rightarrow [0, 1, 0, 1]$$
$$[1, 0] \rightarrow [1, 0, 1, 0]$$
$$[1, 1] \rightarrow [1, 1, 1, 1]$$

Demuestre que este código no es de corrección de error.

80. Suponga que codifica los dígitos binarios 0 y 1 al repetir cada dígito cinco veces. Por tanto,

$$0 \rightarrow [0, 0, 0, 0, 0]$$
$$1 \rightarrow [1, 1, 1, 1, 1]$$

Demuestre que este código puede corregir errores dobles.

¿Cuál es el resultado de codificar los mensajes en los ejercicios 81-83 usando el código de Hamming (7, 4) del ejemplo 3.71?

81. $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ **82.** $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ **83.** $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

Cuando se usa el código de Hamming (7, 4) del ejemplo 3.71, suponga que se reciben los mensajes $\mathbf{c}'$ en los ejercicios 84-86.

Aplique la matriz de control de paridad estándar a $\mathbf{c}'$ *para determinar si ocurrió un error y decodificar correctamente* $\mathbf{c}'$ *para recuperar el vector mensaje original* $\mathbf{x}$.

84. $\mathbf{c}' = [0 \quad 1 \quad 0 \quad 0 \quad 1 \quad 0 \quad 1]^T$

85. $\mathbf{c}' = [1 \quad 1 \quad 0 \quad 0 \quad 1 \quad 1 \quad 0]^T$

86. $\mathbf{c}' = [0 \quad 0 \quad 1 \quad 1 \quad 1 \quad 1 \quad 0]^T$

87. El código de control de paridad en el ejemplo 1.37 es un código $\mathbb{Z}_2^6 \to \mathbb{Z}_2^7$.

(**a**) Encuentre una matriz de control de paridad estándar para este código.

(**b**) Encuentre una matriz generadora estándar.

(**c**) Aplique el Teorema 3.37 para explicar por qué este código no es de corrección de error.

88. Defina un código $\mathbb{Z}_2^2 \to \mathbb{Z}_2^5$ con el uso de una matriz generadora estándar

$$G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

(**a**) Enliste todas las cuatro palabras código.

(**b**) Encuentre la matriz de control de paridad estándar asociada para este código. ¿Este código es de corrección de error (sencillo)?

89. Defina un código $\mathbb{Z}_2^3 \to \mathbb{Z}_2^6$ con la matriz generadora estándar

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

(**a**) Mencione las ocho palabras código.

(**b**) Encuentre la matriz de control de paridad estándar asociada para este código. ¿Este código es de corrección de error (sencillo)?

90. Demuestre que el código del ejemplo 3.70 es un código de Hamming (3, 1).

91. Construya matrices de comprobación de paridad estándar y generadora para un código de Hamming (15, 11).

92. En el Teorema 3.37, demuestre que si $B = A$, entonces $PG\mathbf{x} = \mathbf{0}$ para todo $\mathbf{x}$ en $\mathbb{Z}_2^k$.

93. En el Teorema 3.37, demuestre que, si $\mathbf{p}_i = \mathbf{p}_j$, entonces no es posible determinar si ocurre un error en el componente i-ésimo o j-ésimo del vector recibido.

Repaso del capítulo

Definiciones y conceptos clave

base, 204
Teorema de la base, 208
matriz (vector) columna, 144
espacio columna de una matriz, 201
composición de transformaciones lineales, 225
vector coordenada con respecto a una base, 214
matriz diagonal, 145
dimensión, 209
matriz elemental, 176
teorema fundamental de las matrices invertibles, 178, 212
matriz identidad, 145
inverso de una matriz cuadrada, 169
inverso de una transformación lineal, 227
combinación lineal de matrices, 160
dependencia-independencia lineal de matrices, 163
transformación lineal, 219
factorización *LU*, 187
matriz, 144
suma matricial, 146
factorización matricial, 186
multiplicación matricial, 147
potencias matriciales, 155
negativo de una matriz, 146
espacio nulo de una matriz, 203
nulidad de una matriz, 210
producto exterior, 153
matrices particionadas (multiplicación por bloques), 151, 154
matriz permutación, 193
propiedades del álgebra matricial, 160, 164, 165, 173
rank de una matriz, 210
teorema del rank, 211
representaciones de productos de matrices, 152-154
matriz renglón (vector), 144
espacio renglón de una matriz, 201
múltiplo escalar de una matriz, 146
generador de un conjunto de matrices, 162
matriz cuadrada, 145
matriz estándar de una transformación lineal, 222
subespacio, 198
matriz simétrica, 157
transpuesta de una matriz, 157
matriz cero, 147

Preguntas de repaso

1. Marque cada uno de los siguientes enunciados como verdadero o falso:
   (a) Para cualquier matriz A, tanto AA^T como A^TA están definidas.
   (b) Si A y B son matrices tales que $AB = O$ y $A \neq O$, entonces $B = O$.
   (c) Si A, B y X son matrices invertibles tales que $XA = B$, entonces $X = A^{-1}B$.
   (d) La inversa de una matriz elemental es una matriz elemental.
   (e) La transpuesta de una matriz elemental es una matriz elemental.
   (f) El producto de dos matrices elementales es una matriz elemental.
   (g) Si A es una matriz de $m \times n$, entonces el espacio nulo de A es un subespacio de $\mathbb{R}^n$.
   (h) Todo plano en $\mathbb{R}^3$ es un subespacio bidimensional de $\mathbb{R}^3$.
   (i) La transformación $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida por $T(\mathbf{x}) = -\mathbf{x}$ es una transformación lineal.
   (j) Si $T : \mathbb{R}^4 \rightarrow \mathbb{R}^5$ es una transformación lineal, entonces existe una matriz A de 4×5 tal que $T(\mathbf{x}) = A\mathbf{x}$ para toda $\mathbf{x}$ en el dominio de T.

En los ejercicios 2-7, sea $A = \begin{bmatrix} 1 & 2 \\ 3 & 5 \end{bmatrix}$ *y* $B = \begin{bmatrix} 2 & 0 & -1 \\ 3 & -3 & 4 \end{bmatrix}$.

Calcule la matrices indicadas si es posible.

2. A^2B
3. A^TB^2
4. $B^TA^{-1}B$
5. $(BB^T)^{-1}$
6. $(B^TB)^{-1}$
7. El desarrollo del producto exterior de AA^T
8. Si A es una matriz tal que $A^{-1} = \begin{bmatrix} 1/2 & -1 \\ -3/2 & 4 \end{bmatrix}$, encuentre A.
9. Si $A = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 3 & -1 \\ 0 & 1 & 1 \end{bmatrix}$ y X es una matriz tal que $AX = \begin{bmatrix} -1 & -3 \\ 5 & 0 \\ 3 & -2 \end{bmatrix}$, encuentre X.
10. Si es posible, exprese la matriz $A = \begin{bmatrix} 1 & 2 \\ 4 & 6 \end{bmatrix}$ como un producto de matrices elementales.
11. Si A es una matriz cuadrada tal que $A^3 = O$, demuestre que $(I - A)^{-1} = I + A + A^2$.
12. Encuentre una factorización LU de $A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}$.
13. Encuentre bases para el espacio renglón, espacio columna y espacio nulo de $A = \begin{bmatrix} 2 & -4 & 5 & 8 & 5 \\ 1 & -2 & 2 & 3 & 1 \\ 4 & -8 & 3 & 2 & 6 \end{bmatrix}$.
14. Suponga que las matrices A y B son equivalentes en renglón. ¿Tienen el mismo espacio renglón? ¿Por qué sí o por qué no? ¿A y B tienen el mismo espacio columna? ¿Por qué sí o por qué no?
15. Si A es una matriz invertible, explique por qué A y A^T deben tener el mismo espacio nulo. ¿Esto es verdad si A es una matriz cuadrada no invertible? Explique.
16. Si A es una matriz cuadrada cuyos renglones suman el vector cero, explique por qué A no puede ser invertible.
17. Sea A una matriz de $m \times n$ con columnas linealmente independientes. Explique por qué A^TA debe ser una matriz invertible. ¿AA^T también debe ser invertible? Explique.
18. Encuentre una transformación lineal $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ tal que $T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $T\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}$.
19. Encuentre la matriz estándar de la transformación lineal $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ que corresponda a una rotación de 45° en sentido contrario al de las manecillas del reloj en torno al origen seguida por una proyección sobre la recta $y = -2x$.
20. Suponga que $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una transformación lineal y suponga que $\mathbf{v}$ es un vector tal que $T(\mathbf{v}) \neq \mathbf{0}$, pero $T^2(\mathbf{v}) = \mathbf{0}$ (donde $T^2 = T \circ T$). Demuestre que $\mathbf{v}$ y $T(\mathbf{v})$ son linealmente independientes.

4 Eigenvalores y eigenvectores

En la literatura se ha usado casi toda combinación de los adjetivos propio, latente, característico, eigen y secular, con los sustantivos raíz, número y valor, para lo que llamamos un valor propio.

—Paul R. Halmos
Finite Dimensional Vector Spaces
(2a. edición), Van Nostrand, 1958, p. 102

4.0 Introducción: un sistema dinámico de grafos

CAS

En el último capítulo se vio que iterar la multiplicación de matrices con frecuencia produce resultados interesantes. Tanto en las cadenas de Markov como en el modelo de Leslie de crecimiento poblacional se muestran estados estacionarios en ciertas situaciones. Una de las metas de este capítulo es ayudarle a comprender tal comportamiento. Primero se observará un proceso iterativo, o ***sistema dinámico***, que usa matrices. (En los problemas siguientes encontrará útil usar un CAS o una calculadora con capacidades matriciales para facilitar los cálculos.)

El ejemplo involucra grafos (vea la sección 3.7). Un ***grafo completo*** es aquel en el que todo vértice es adyacente a todo otro vértice. Si un grafo completo tiene n vértices, se denota K_n. Por ejemplo, la figura 4.1 muestra una representación de K_4.

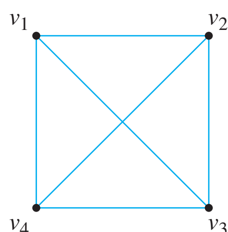


Figura 4.1
K_4

Problema 1 Escoja cualquier vector $\mathbf{x}$ en $\mathbb{R}^4$ con entradas no negativas y marque los vértices de K_4 con los componentes de $\mathbf{x}$, de modo que v_1 se etiquete con x_1, etcétera. Calcule la matriz de adyacencia A de K_4 y vuelva a etiquetar los vértices del grafo con los correspondientes componentes de $A\mathbf{x}$. Intente esto para varios vectores $\mathbf{x}$ y explique, en términos del grafo, cómo pueden determinarse las nuevas etiquetas a partir de las etiquetas anteriores.

Problema 2 Ahora *itere* el proceso en el problema 1. Esto es, para una elección dada de $\mathbf{x}$, vuelva a etiquetar los vértices como se describió anteriormente y luego aplique de nuevo A (y de nuevo, y de nuevo) hasta que surja un patrón. Dado que los componentes de los vectores se volverán bastante grandes, se les reducirá a *escala* al dividir cada vector entre su *componente más grande* después de cada iteración. Por tanto, si un cálculo resulta en el vector

$$\begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \end{bmatrix}$$

se le sustituirá con

$$\frac{1}{4}\begin{bmatrix} 4 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.5 \\ 0.25 \\ 0.25 \end{bmatrix}$$

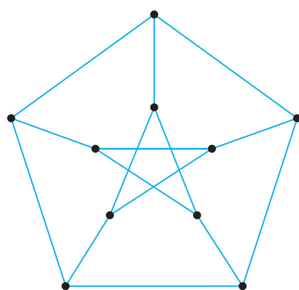


Figura 4.2

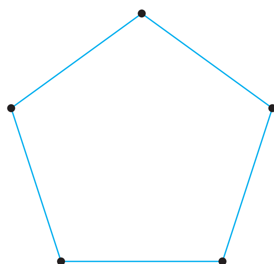


Figura 4.3

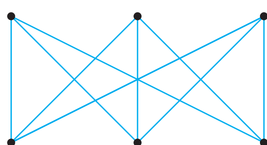


Figura 4.4

Note que este proceso garantiza que el componente más grande de cada vector ahora será 1. Haga esto para K_4, luego K_3 y K_5. Use al menos diez iteraciones y precisión a dos lugares decimales. ¿Qué parece ocurrir?

Problema 3 Debe haber notado que, en cada caso, el vector que etiqueta tiende a cierto valor (¡una etiqueta de estado estacionario!). Etiquete los vértices de los grafos completos con este vector de estado estacionario y aplique la matriz de adyacencia A una vez más (sin escala). ¿Cuál es la relación entre las nuevas etiquetas y las anteriores?

Problema 4 Haga una conjetura acerca del caso general K_n. ¿Cuál es la etiqueta de estado estacionario? ¿Qué ocurre si se etiqueta K_n con el vector de estado estacionario y se aplica la matriz de adyacencia A sin escala?

Problema 5 En la figura 4.2 se muestra el ***grafo de Petersen***. Repita el proceso de los problemas 1 al 3 con este grafo.

Ahora se explorará el proceso con alguna otra clase de grafos para ver si se comportan en esta forma. El ***ciclo*** C_n es el grafo con n vértices ordenados en forma cíclica. Por ejemplo, C_5 es el grafo que se muestra en la figura 4.3.

Problema 6 Repita el proceso en los problemas 1 a 3, con ciclos C_n para varios valores *impares* de n y haga una conjetura acerca del caso general.

Problema 7 Repita el problema 6 con valores *pares* de n. ¿Qué sucede?

Un grafo bipartito es un ***grafo bipartito completo*** (vea los ejercicios 74-78 de la sección 3.7) si sus vértices pueden partirse en conjuntos U y V tales que cada vértice en U es adyacente a cada vértice en V y viceversa. Si U y V tienen cada uno n vértices, entonces el grafo se denota $K_{n,n}$. Por ejemplo, $K_{3,3}$ es el grafo de la figura 4.4.

Problema 8 Repita el proceso de los problemas 1 a 3 con grafos bipartitos completos $K_{n,n}$ para varios valores de n. ¿Qué sucede?

Hacia el final de este capítulo, estará en posición de explicar las observaciones realizadas en esta introducción.

4.1 Introducción a eigenvalores y eigenvectores

En el capítulo 3 se encontró la noción de vector de estado estacionario en el contexto de dos aplicaciones: cadenas de Markov y el modelo de Leslie de crecimiento poblacional. Para una cadena de Markov con matriz de transición P, un vector de estado estacionario $\mathbf{x}$ tiene la propiedad de que $P\mathbf{x} = \mathbf{x}$; para una matriz de Leslie L, un vector de estado estacionario era un vector de población $\mathbf{x}$ que satisfacía $L\mathbf{x} = r\mathbf{x}$, donde r representaba la tasa de crecimiento de estado estacionario. Por ejemplo, se vio que

$$\begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}\begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 & 4 & 3 \\ 0.5 & 0 & 0 \\ 0 & 0.25 & 0 \end{bmatrix}\begin{bmatrix} 18 \\ 6 \\ 1 \end{bmatrix} = 1.5\begin{bmatrix} 18 \\ 6 \\ 1 \end{bmatrix}$$

El adjetivo alemán *eigen* significa "propio" o "característico de". *Eigenvalores* y *eigenvectores* son características de una matriz en el sentido de que contienen información importante acerca de la naturaleza de la matriz. La letra λ (lambda), la equivalente griega de la letra L, se usa para eigenvalores porque en una época también se conocían como *valores latentes*. El comentario es para los que hablan inglés.

En este capítulo se investiga este fenómeno de manera más general. Esto es, para una matriz cuadrada A, se pregunta si existen vectores distintos de cero $\mathbf{x}$ tales que $A\mathbf{x}$ es justo un múltiplo escalar de $\mathbf{x}$. Este es el *problema del eigenvalor* y es uno de los problemas centrales en álgebra lineal. Tiene aplicaciones en todas las matemáticas y también en muchos otros campos.

Definición Sea A una matriz de $n \times n$. Un escalar λ se llama ***eigenvalor*** de A si existe un vector $\mathbf{x}$ distinto de cero tal que $A\mathbf{x} = \lambda\mathbf{x}$. Tal vector $\mathbf{x}$ se llama ***eigenvector*** de A correspondiente a λ.

Ejemplo 4.1

Demuestre que $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ es un eigenvector de $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ y encuentre el eigenvalor correspondiente.

Solución Calcule

$$A\mathbf{x} = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 4\mathbf{x}$$

de donde se tiene que $\mathbf{x}$ es un eigenvector de A que corresponde al eigenvalor 4.

Ejemplo 4.2

Demuestre que 5 es un eigenvalor de $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$ y determine todos los eigenvectores correspondientes a este eigenvalor.

Solución Debe demostrar que existe un vector $\mathbf{x}$ *distinto de cero* tal que $A\mathbf{x} = 5\mathbf{x}$. Pero esta ecuación es equivalente a la ecuación $(A - 5I)\mathbf{x} = \mathbf{0}$, así que es necesario calcular el espacio nulo de la matriz $A - 5I$. Se descubre que

$$A - 5I = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} - \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} -4 & 2 \\ 4 & -2 \end{bmatrix}$$

Dado que las columnas de esta matriz es claro que son linealmente dependientes, el teorema fundamental de las matrices invertibles implica que su espacio nulo es distinto de cero. Por tanto, $A\mathbf{x} = 5\mathbf{x}$ tiene una solución no trivial, así que 5 es un eigenvalor de A. Sus eigenvectores se encuentran al calcular el espacio nulo:

$$[A - 5I \mid \mathbf{0}] = \left[\begin{array}{rr|r} -4 & 2 & 0 \\ 4 & -2 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{array}\right]$$

Por tanto, si $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ es un eigenvector que corresponde al eigenvalor 5, satisface $x_1 - \frac{1}{2}x_2 = 0$ o $x_1 = \frac{1}{2}x_2$, de modo que dichos eigenvectores son de la forma

$$\mathbf{x} = \begin{bmatrix} \frac{1}{2}x_2 \\ x_2 \end{bmatrix}$$

Esto es, son los múltiplos distintos de cero de $\begin{bmatrix} \frac{1}{2} \\ 1 \end{bmatrix}$ (o, de manera equivalente, los múltiplos distintos de cero de $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$).

El conjunto de todos los eigenvectores correspondientes a un eigenvalor λ de una matriz A de $n \times n$ es justo el conjunto de vectores *distintos de cero* en el espacio nulo de $A - \lambda I$. Por tanto, este conjunto de eigenvectores, junto con el vector cero en ^{n}, *es* el espacio nulo de $A - \lambda I$.

Definición Sea A una matriz de $n \times n$ y sea λ un eigenvalor de A. La colección de todos los eigenvectores correspondientes a λ, junto con el vector cero, se llama ***eigenespacio*** de λ y se denota E_λ.

Por tanto, en el ejemplo 4.2, $E_5 = \left\{ t\begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$.

Ejemplo 4.3

Demuestre que $\lambda = 6$ es un eigenvalor de $A = \begin{bmatrix} 7 & 1 & -2 \\ -3 & 3 & 6 \\ 2 & 2 & 2 \end{bmatrix}$ y encuentre una base para su eigenespacio.

Solución Como en el ejemplo 4.2, calcule el espacio nulo de $A - 6I$. La reducción por renglones produce

$$A - 6I = \begin{bmatrix} 1 & 1 & -2 \\ -3 & -3 & 6 \\ 2 & 2 & -4 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

de donde se ve que el espacio nulo de $A - 6I$ es distinto de cero. Por tanto, 6 es un eigenvalor de A y los eigenvectores correspondientes a este eigenvalor satisfacen $x_1 + x_2 - 2x_3 = 0$ o $x_1 = -x_2 + 2x_3$. Se tiene que

$$E_6 = \left\{ \begin{bmatrix} -x_2 + 2x_3 \\ x_2 \\ x_3 \end{bmatrix} \right\} = \left\{ x_2\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\} = \text{gen}\left(\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right)$$

En $\mathbb{R}^2$ puede darse una interpretación geométrica de la noción de un eigenvector. La ecuación $A\mathbf{x} = \lambda\mathbf{x}$ dice que los vectores $A\mathbf{x}$ y $\mathbf{x}$ son paralelos. Por tanto, $\mathbf{x}$ es un eigenvector de A si y sólo si A transforma $\mathbf{x}$ en un vector paralelo [o, de manera equivalente, si y sólo si $T_A(\mathbf{x})$ es paralela a $\mathbf{x}$, donde T_A es la transformación matricial correspondiente a A].

Ejemplo 4.4

Encuentre geométricamente los eigenvectores y eigenvalores de $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$.

Solución Se reconoce que A es la matriz de una reflexión F en el eje x (vea el ejemplo 3.56). Los únicos vectores que F mapea paralelos a ellos mismos son vectores paralelos al eje y (es decir: múltiplos de $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$) que están invertidos (eigenvalor -1), y vectores paralelos al eje x (es decir, múltiplos de $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$) que se envían a ellos mismos (eigenvalor 1) (vea la figura 4.5). En concordancia, $\lambda = -1$ y $\lambda = 1$ son los eigenvalores de A, y los correspondientes eigenespacios son

$$E_{-1} = \text{gen}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) \quad \text{y} \quad E_1 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$$

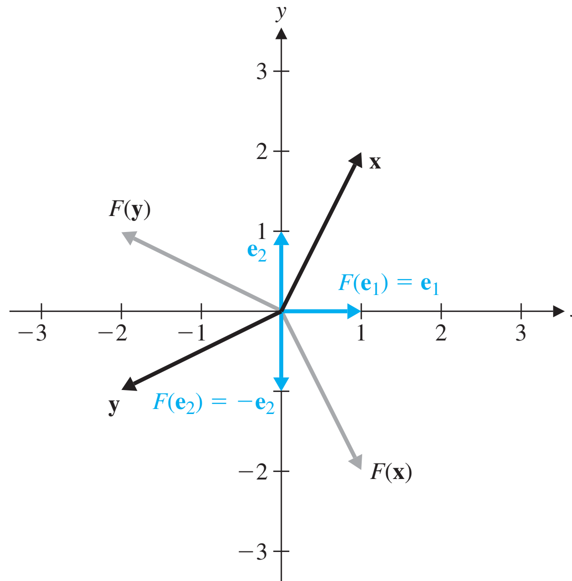


Figura 4.5
Los eigenvectores de una reflexión

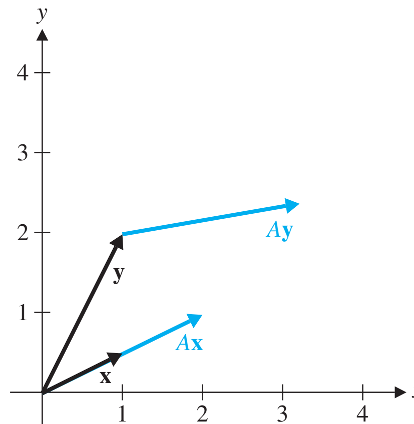


Figura 4.6

La discusión se basa en el artículo "Eigenpictures: Picturing the Eigenvector Problem" de Steven Schonefeld en *The College Mathematics Journal 26* (1996), pp. 316-319.

Otra forma de pensar geométricamente en los eigenvectores es dibujar $\mathbf{x}$ y $A\mathbf{x}$ de origen a punta. Entonces $\mathbf{x}$ será un eigenvector de A si y sólo si $\mathbf{x}$ y $A\mathbf{x}$ están alineados en una recta. En la figura 4.6, $\mathbf{x}$ es un eigenvector de A pero $\mathbf{y}$ no lo es.

Si $\mathbf{x}$ es un eigenvector de A correspondiente al eigenvalor λ, entonces también lo es cualquier múltiplo de $\mathbf{x}$ distinto de cero. De este modo, si quiere encontrar eigenvectores geométricamente, sólo es necesario considerar el efecto de A sobre vectores *unitarios*. La figura 4.7(a) muestra lo que ocurre cuando se transforman vectores unitarios con la matriz $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ del ejemplo 4.1 y muestra los resultados origen a punta, como en la figura 4.6. Puede ver que el vector $\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ es un eigenvector, pero note también que parece haber un eigenvector en el segundo cuadrante. De hecho, este es el caso, y evidencia ser el vector $\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

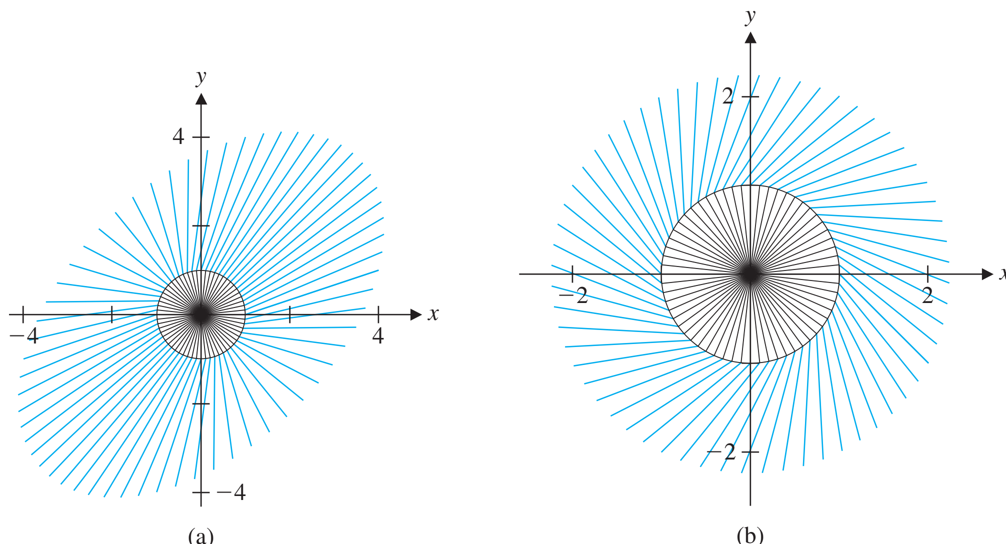


Figura 4.7

En la figura 4.7(b) se ve lo que ocurre cuando se usa la matriz $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$. ¡No hay eigenvectores!

Ahora sabe cómo encontrar eigenvectores una vez que tiene los eigenvalores correspondientes, y tiene una interpretación geométrica de ellos, pero surge una pregunta: ¿cómo se encuentran por primera vez los eigenvalores de una matriz dada? La clave es la observación de que λ es un eigenvalor de A si y sólo si el espacio nulo de $A - \lambda I$ es no trivial.

Recuerde de la sección 3.3 que el determinante de una matriz de 2×2 $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es la expresión $A = ad - bc$, y A es invertible si y sólo si det A es distinta de cero. Más aún, el teorema fundamental de las matrices invertibles garantiza que una matriz tiene un espacio nulo no trivial si y sólo si no es invertible; es decir, si y sólo si su determinante es cero. Al juntar estos hechos, se ve que (al menos para matrices de 2×2) λ es un eigenvalor de A si y sólo si $\det(A - \lambda I) = 0$. Este hecho caracteriza a los eigenvalores, y pronto se le generalizará a matrices cuadradas de tamaño arbitrario. Sin embargo, por el momento, vea cómo usarlo con matrices de 2×2.

Ejemplo 4.5

Encuentre todos los eigenvalores y los correspondientes eigenvectores de la matriz $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ del ejemplo 4.1.

Solución Las observaciones anteriores muestran que debe encontrar todas las soluciones λ de la ecuación $\det(A - \lambda I) = 0$. Puesto que

$$\det(A - \lambda I) = \det\begin{bmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix} = (3-\lambda)(3-\lambda) - 1 = \lambda^2 - 6\lambda + 8$$

es necesario resolver la ecuación cuadrática $\lambda^2 - 6\lambda + 8 = 0$. Se encuentra fácilmente que las soluciones a esta ecuación son $\lambda = 4$ y $\lambda = 2$. Por tanto, esos son los eigenvalores de A.

Para encontrar los eigenvectores correspondientes al eigenvalor $\lambda = 4$, calcule el espacio nulo de $A - 4I$. Se encuentra que

$$[A - 4I \mid \mathbf{0}] = \left[\begin{array}{rr|r} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

de donde se tiene que $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ es un eigenvector que corresponde a $\lambda = 4$ si y sólo si $x_1 - x_2 = 0$ o $x_1 = x_2$. Por tanto, el eigenespacio $E_4 = \left\{ \begin{bmatrix} x_2 \\ x_2 \end{bmatrix} \right\} = \left\{ x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)$.

De igual modo, para $\lambda = 2$, se tiene

$$[A - 2I \mid \mathbf{0}] = \left[\begin{array}{rr|r} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right] \rightarrow \left[\begin{array}{rr|r} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

de modo que $\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ es un eigenvector que corresponde a $\lambda = 2$ si y sólo si $y_1 + y_2 = 0$ o $y_1 = -y_2$. Por tanto, el eigenespacio $E_2 = \left\{ \begin{bmatrix} -y_2 \\ y_2 \end{bmatrix} \right\} = \left\{ y_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right\} = \text{gen}\left(\begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$.

La figura 4.8 muestra gráficamente cómo se transformaron los eigenvectores de A cuando se multiplicaron por A: un eigenvector $\mathbf{x}$ en el eigenespacio E_4 se transforma en $4\mathbf{x}$, y un eigenvector $\mathbf{y}$ en el eigenespacio E_2 se transforma en $2\mathbf{y}$. Como muestra la figura 4.7(a), los eigenvectores de A son los *únicos* vectores en $\mathbb{R}^2$ que se transforman en múltiplos escalares de ellos mismos cuando se multiplican por A.

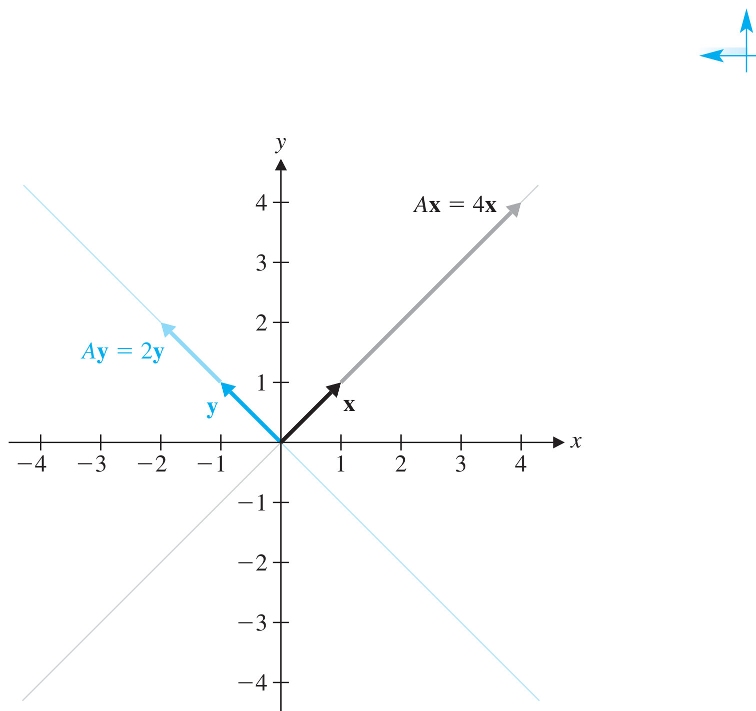


Figura 4.8
Cómo A transforma eigenvectores

Comentario Recordará que una ecuación polinomial con coeficientes reales (como la ecuación cuadrática del ejemplo 4.5) no necesita tener raíces reales; puede tener raíces *complejas*. (Vea el Apéndice C.) También es posible calcular eigenvalores y eigenvectores cuando las entradas de una matriz provienen de $\mathbb{Z}_p$, donde p es primo. Por tanto, es importante especificar el escenario en el que se intenta trabajar antes de calcular los eigenvalores de una matriz. Sin embargo, a menos que se especifique de otro modo, los eigenvalores de una matriz cuyas entradas sean números *reales* se supondrá que también son reales.

Ejemplo 4.6 Interprete la matriz del ejemplo 4.5 como una matriz sobre $\mathbb{Z}_3$ y encuentre sus eigenvalores en dicho campo.

Solución La solución procede exactamente como antes, excepto que se trabaja en módulo 3. Por tanto, la ecuación cuadrática $\lambda^2 - 6\lambda + 8 = 0$ se convierte en $\lambda^2 + 2 = 0$. Esta ecuación es la misma que $\lambda^2 = -2 = 1$, lo que produce $\lambda = 1$ y $\lambda = -1 = 2$ como los eigenvalores en $\mathbb{Z}_3$. (Compruebe que obtendría la misma respuesta al *primero* reducir A a módulo 3 para obtener $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ y luego trabajar con esta matriz.)

Ejemplo 4.7 Encuentre los eigenvalores de $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (a) sobre $\mathbb{R}$ y (b) sobre los números complejos $\mathbb{C}$.

Solución Debe resolver la ecuación

$$0 = \det(A - \lambda I) = \det\begin{bmatrix} -\lambda & -1 \\ 1 & -\lambda \end{bmatrix} = \lambda^2 + 1$$

a + bi

(a) Sobre $\mathbb{R}$ no hay soluciones, de modo que A no tiene eigenvalores reales.

(b) Sobre $\mathbb{C}$ las soluciones son $\lambda = i$ y $\lambda = -i$. (Vea el Apéndice C.)

En la siguiente sección se extenderá la noción de determinante de 2×2 a matrices de $n \times n$, que a su vez permitirán encontrar los eigenvalores de matrices cuadradas arbitrarias. (De hecho, esto no es totalmente cierto, pero al menos podrá encontrar una ecuación polinomial que deben satisfacer los eigenvalores de una matriz dada.)

Ejercicios 4.1

En los ejercicios 1-6, demuestre que **v** *es un eigenvector de A y encuentre el eigenvalor correspondiente.*

1. $A = \begin{bmatrix} 0 & 3 \\ 3 & 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

2. $A = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$

3. $A = \begin{bmatrix} -1 & 1 \\ 6 & 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

4. $A = \begin{bmatrix} 4 & -2 \\ 5 & -7 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 10 \end{bmatrix}$

5. $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 3 & -1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 7-12, demuestre que λ es un eigenvalor de A y encuentre un eigenvector correspondiente a este eigenvalor.

7. $A = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}, \lambda = 3$

8. $A = \begin{bmatrix} 2 & 2 \\ 2 & -1 \end{bmatrix}, \lambda = -2$

9. $A = \begin{bmatrix} 0 & 4 \\ -1 & 5 \end{bmatrix}, \lambda = 1$

10. $A = \begin{bmatrix} 0 & 4 \\ -1 & 5 \end{bmatrix}, \lambda = 4$

11. $A = \begin{bmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}, \lambda = -1$

12. $A = \begin{bmatrix} 3 & 1 & -1 \\ 0 & 1 & 2 \\ 4 & 2 & 0 \end{bmatrix}, \lambda = 3$

En los ejercicios 13-18, encuentre geométricamente los eigenvalores y eigenvectores de A.

13. $A = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ (reflexión en el eje y)

14. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ (reflexión en la recta $y = x$)

15. $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ (proyección sobre el eje x)

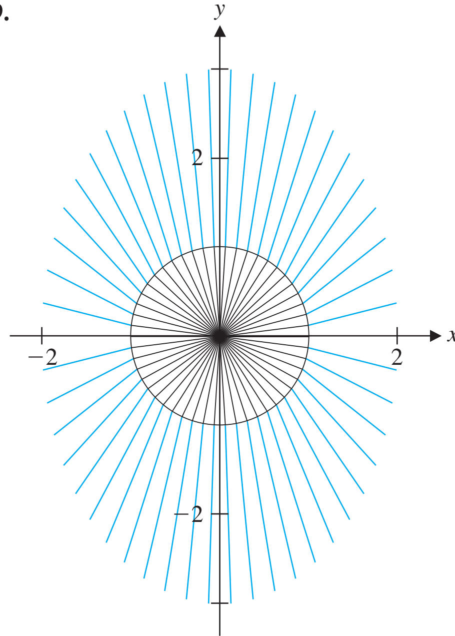
16. $A = \begin{bmatrix} \frac{16}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{9}{25} \end{bmatrix}$ (proyección sobre la recta que pasa por el origen con vector director $\begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$)

17. $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ (estiramiento por un factor de 2 horizontalmente y un factor de 3 verticalmente)

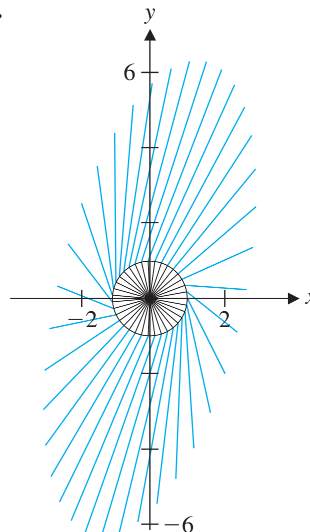
18. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ (rotación de 90° en sentido contrario al de las manecillas del reloj en torno al origen)

En los ejercicios 19-22, los vectores unitarios **x** *en* $\mathbb{R}^2$ *y sus imágenes A***x** *bajo la acción de una matriz A de 2 × 2 se dibujan origen a punta, como en la figura 4.7. Estime los eigenvectores y eigenvalores de A a partir de cada "eigendibujo".*

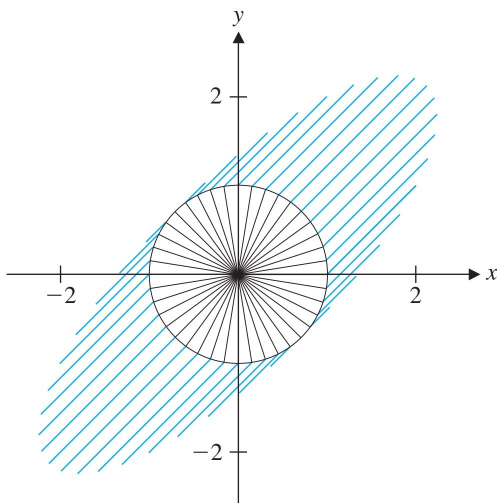
19.



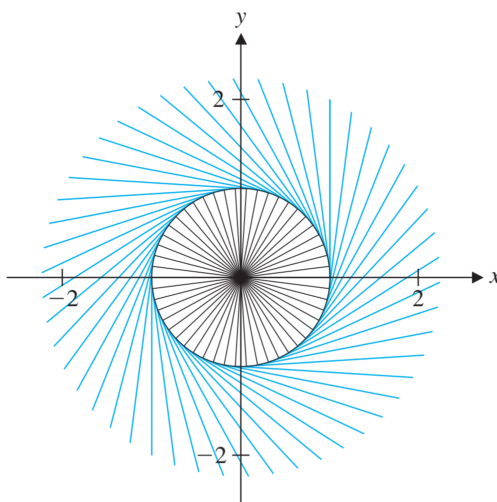
20.



21.



22.



En los ejercicios 23-26, use el método del ejemplo 4.5 para encontrar todos los eigenvalores de la matriz A. Obtenga bases para cada uno de los eigenespacios correspondientes. Ilustre los eigenespacios y el efecto de multiplicar los eigenvectores por A, como en la figura 4.8.

23. $A = \begin{bmatrix} 4 & -1 \\ 2 & 1 \end{bmatrix}$ **24.** $A = \begin{bmatrix} 0 & 2 \\ 8 & 6 \end{bmatrix}$

25. $A = \begin{bmatrix} 2 & 5 \\ 0 & 2 \end{bmatrix}$ **26.** $A = \begin{bmatrix} 2 & 1 \\ -1 & 2 \end{bmatrix}$

a + bi

En los ejercicios 27-30, encuentre todos los eigenvalores de la matriz A en los números complejos $\mathbb{C}$. Obtenga bases para cada uno de los eigenespacios correspondientes.

27. $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ **28.** $A = \begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix}$

29. $A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$ **30.** $A = \begin{bmatrix} 4 & 1-2i \\ 1+2i & 0 \end{bmatrix}$

En los ejercicios 31-34, encuentre todos los eigenvalores de la matriz A en el $\mathbb{Z}_p$ indicado.

31. $A = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}$ en $\mathbb{Z}_3$ **32.** $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ en $\mathbb{Z}_3$

33. $A = \begin{bmatrix} 3 & 1 \\ 4 & 0 \end{bmatrix}$ en $\mathbb{Z}_5$ **34.** $A = \begin{bmatrix} 1 & 4 \\ 4 & 0 \end{bmatrix}$ en $\mathbb{Z}_5$

35. **(a)** Demuestre que los eigenvalores de la matriz 2×2

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

son soluciones de la ecuación cuadrática $\lambda^2 - \operatorname{tr}(A)\lambda + \det A = 0$, donde $\operatorname{tr}(A)$ es la traza de A. (Vea la página 168.)

(b) Demuestre que los eigenvalores de la matriz A en el inciso (a) son

$$\lambda = \tfrac{1}{2}\left(a + d \pm \sqrt{(a-d)^2 + 4bc}\right)$$

(c) Demuestre que la traza y el determinante de la matriz A en el inciso (a) están dados por

$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 \quad \text{y} \quad \det A = \lambda_1 \lambda_2$$

donde λ_1 y λ_2 son los eigenvalores de A.

36. Considere nuevamente la matriz A en el ejercicio 35. Proporcione las condiciones para a, b, c y d tales que A tenga

(a) dos eigenvalores reales distintos,

(b) un eigenvalor real y

(c) ningún eigenvalor real.

37. Demuestre que los eigenvalores de la matriz triangular superior

$$A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

son $\lambda = a$ y $\lambda = d$, y encuentre los correspondientes eigenespacios.

a + bi **38.** Sean a y b números reales. Encuentre los eigenvalores y eigenespacios correspondientes de

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

en los números complejos.

4.2 Determinantes

Históricamente, los determinantes precedían a las matrices, un hecho curioso a la luz de la forma como se enseña el álgebra lineal en la actualidad, con las matrices antes que los determinantes. No obstante, los determinantes surgen independientemente de las matrices en la solución de muchos problemas prácticos, y la teoría de los determinantes estaba bien desarrollada casi dos siglos antes de que a la matrices se les otorgara un valor de estudio como tal. Al final de esta sección se presenta una reseña histórica de los determinantes.

Recuerde que el determinante de la matriz $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ de 2×2 es

$$\det A = a_{11}a_{22} - a_{12}a_{21}$$

Esta expresión se encontró por primera vez cuando se determinaron formas para calcular la inversa de una matriz. En particular, se encontró que

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

El determinante de una matriz A en ocasiones también se denota por $|A|$, de modo que para la matriz $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ de 2×2 es posible escribir

$$|A| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Advertencia Esta notación para el determinante recuerda la notación del valor absoluto. Es fácil confundir la notación $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, para determinante, con la notación $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, para la matriz en sí. No las confunda. Por fortuna, usualmente será claro por el contexto lo que se pretende.

Se define el determinante de una matriz $A = [a]$ de 1×1 como

$$\det A = |a| = a$$

(Note que realmente debe tener cuidado con esta notación: en este caso, $|a|$ *no* denota el valor absoluto de a.) ¿Cómo se define entonces el determinante de una matriz de 3×3? Si a su CAS le solicita la inversa de

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

la respuesta será equivalente a

$$A^{-1} = \frac{1}{\Delta} \begin{bmatrix} ei - fh & ch - bi & bf - ce \\ fg - di & ai - cg & cd - af \\ dh - eg & bg - ah & ae - bd \end{bmatrix}$$

donde $\Delta = aei - afh - bdi + bfg + cdh - ceg$. Observe que

$$\begin{aligned} \Delta &= aei - afh - bdi + bfg + cdh - ceg \\ &= a(ei - fh) - b(di - fg) + c(dh - eg) \\ &= a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix} \end{aligned}$$

y que cada una de las entradas en la porción de matriz de A^{-1} parezca ser el determinante de una submatriz de 2×2 de A. De hecho, esto es cierto, y es la base de la definición del determinante de una matriz de 3×3. La definición es *recursiva* en el sentido de que el determinante de una matriz de 3×3 se define en términos de determinantes de matrices de 2×2.

Definición Sea $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$. Entonces el ***determinante*** de A es el escalar

$$\det A = |A| = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \qquad (1)$$

Note que cada uno de los determinantes de 2×2 se obtiene al borrar el renglón y la columna de A que contienen la entrada por la que se multiplica el determinante. Por ejemplo, el primer sumando es a_{11} multiplicado por el determinante de la submatriz obtenida al borrar el renglón 1 y la columna 1. Note también que los signos más y menos se alternan en la ecuación (1). Si se denota con A_{ij} la submatriz de una matriz A obtenida al borrar el renglón i y la columna j, entonces la ecuación (1) se puede abreviar como

$$\begin{aligned} \det A &= a_{11} \det A_{11} - a_{12} \det A_{12} + a_{13} \det A_{13} \\ &= \sum_{j=1}^{3} (-1)^{1+j} a_{1j} \det A_{1j} \end{aligned}$$

Para cualquier matriz cuadrada A, $\det A_{ij}$ se llama ***menor-(i, j)*** de A.

Ejemplo 4.8 Calcule el determinante de

$$A = \begin{bmatrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{bmatrix}$$

Solución Calcule

$$\begin{aligned} \det A &= 5\begin{vmatrix} 0 & 2 \\ -1 & 3 \end{vmatrix} - (-3)\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + 2\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} \\ &= 5(0 - (-2)) + 3(3 - 4) + 2(-1 - 0) \\ &= 5(2) + 3(-1) + 2(-1) = 5 \end{aligned}$$

Con un poco de práctica, debe descubrir que puede trabajar fácilmente determinantes de 2×2 en su cabeza. Entonces, escribir la segunda línea en la solución anterior es innecesario.

Otro método para calcular el determinante de una matriz de 3×3 es análogo al método para calcular el determinante de una matriz de 2×2. Copie las primeras dos columnas de A a la derecha de la matriz y tome los productos de los elementos en las seis

diagonales que se muestran a continuación. Agregue signos más a los productos de las diagonales con pendiente descendente y signos menos a los productos de las diagonales con pendiente ascendente.

$$\begin{matrix} & & & - & - & - \\ \left[\begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix}\right] & \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix} \\ & & & + & + & + \end{matrix} \tag{2}$$

Este método produce

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

En el ejercicio 19 se le pedirá comprobar que este resultado está de acuerdo con el de la ecuación (1) para un determinante de 3 × 3.

Ejemplo 4.9

Calcule el determinante de la matriz en el ejemplo 4.8 con el método que se muestra en (2).

Solución Agregue a A sus primeras dos columnas y calcule los seis productos indicados:

$$\begin{matrix} & & & 0 & -10 & -9 \\ \left[\begin{matrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{matrix}\right] & \begin{matrix} 5 & -3 \\ 1 & 0 \\ 2 & -1 \end{matrix} \\ & & & 0 & -12 & -2 \end{matrix}$$

Sumar los tres productos en la parte baja y restar los tres productos en la parte superior produce

$$\det A = 0 + (-12) + (-2) - 0 - (-10) - (-9) = 5$$

como antes.

Advertencia Está a punto de definir determinantes para matrices cuadradas arbitrarias. Sin embargo, *no* hay un análogo del método en el ejemplo 4.9 para matrices más grandes. *Sólo* es válido para matrices de 3 × 3.

Determinantes de matrices de $n \times n$

La definición del determinante de una matriz de 3 × 3 se extiende naturalmente a matrices cuadradas arbitrarias.

Definición Sea $A = [a_{ij}]$ una matriz de $n \times n$, donde $n \geq 2$. Entonces el ***determinante*** de A es el escalar

$$\begin{aligned} \det A = |A| &= a_{11} \det A_{11} - a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n} \\ &= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det A_{1j} \end{aligned} \tag{3}$$

Es conveniente combinar un menor con su signo más o menos. Para este fin, se define el ***cofactor-(i, j) de A*** como

$$C_{ij} = (-1)^{i+j} \det A_{ij}$$

Con esta notación, la definición (3) se convierte en

$$\det A = \sum_{j=1}^{n} a_{1j} C_{1j} \tag{4}$$

El ejercicio 20 le pide comprobar que esta definición proporciona correctamente la fórmula para el determinante de una matriz de 2×2 cuando $n = 2$.

Con frecuencia, a la definición (4) se le conoce como ***expansión por cofactores a lo largo del primer renglón***. ¡Es un hecho sorprendente que se obtenga exactamente el mismo resultado al expandir a lo largo de *cualquier renglón* (o incluso *cualquier columna*)! Este hecho se resume como teorema, pero la demostración se difiere hasta el final de esta sección (pues es un poco larga e interrumpiría el análisis si se presentara aquí).

Teorema 4.1

Teorema de expansión de Laplace

El determinante de una matriz A de $n \times n$, donde $n \geq 2$, puede calcularse como

$$\begin{aligned} \det A &= a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} \\ &= \sum_{j=1}^{n} a_{ij}C_{ij} \end{aligned} \tag{5}$$

(que es la ***expansión por cofactores a lo largo del i-ésimo renglón***) y también como

$$\begin{aligned} \det A &= a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} \\ &= \sum_{i=1}^{n} a_{ij}C_{ij} \end{aligned} \tag{6}$$

(la ***expansión por cofactores a lo largo de la j-ésima columna***).

Dado que $C_{ij} = (-1)^{i+j} \det A_{ij}$, cada cofactor es más o menos el correspondiente menor, donde $(-1)^{i+j}$ proporciona el signo correcto. Una forma rápida de determinar si el signo es $+$ o $-$ es recordar que los signos forman un patrón de "tablero de ajedrez":

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Ejemplo 4.10

Calcule el determinante de la matriz

$$A = \begin{bmatrix} 5 & -3 & 2 \\ 1 & 0 & 2 \\ 2 & -1 & 3 \end{bmatrix}$$

por (a) expansión por cofactores a lo largo del tercer renglón y (b) expansión por cofactores a lo largo de la segunda columna.

Solución

(a) Calcule

$$\begin{aligned} \det A &= a_{31}C_{31} + a_{32}C_{32} + a_{33}C_{33} \\ &= 2\begin{vmatrix} -3 & 2 \\ 0 & 2 \end{vmatrix} - (-1)\begin{vmatrix} 5 & 2 \\ 1 & 2 \end{vmatrix} + 3\begin{vmatrix} 5 & -3 \\ 1 & 0 \end{vmatrix} \\ &= 2(-6) + 8 + 3(3) \\ &= 5 \end{aligned}$$

(b) En este caso se tiene

$$\begin{aligned} \det A &= a_{12}C_{12} + a_{22}C_{22} + a_{32}C_{32} \\ &= -(-3)\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} + 0\begin{vmatrix} 5 & 2 \\ 2 & 3 \end{vmatrix} - (-1)\begin{vmatrix} 5 & 2 \\ 1 & 2 \end{vmatrix} \\ &= 3(-1) + 0 + 8 \\ &= 5 \end{aligned}$$

Pierre Simon Laplace (1749–1827) nació en Normandía, Francia, y se esperaba que se convirtiera en clérigo hasta que se descubrió su talento matemático en la escuela. Hizo muchas importantes aportaciones al cálculo, la probabilidad y la astronomía. Fue examinador del joven Napoleón Bonaparte en el Real Cuerpo de Artilleros y, más tarde, cuando Napoleón estuvo en el poder, sirvió brevemente como ministro del Interior y luego como canciller del Senado. Laplace fue honrado con el título de conde del Imperio en 1806 y en 1817 recibió el título de marqués de Laplace.

Note que en el inciso (b) del ejemplo 4.10 fue necesario hacer menos cálculos que en el inciso (a) porque se expandió a lo largo de una columna que contenía una entrada cero: a saber, a_{22}; por tanto, no fue necesario calcular C_{22}. Se tiene que el teorema de expansión de Laplace es más útil cuando la matriz contiene un renglón o una columna con muchos ceros, pues al elegir expandir a lo largo de dicho renglón o columna, se minimiza el número de cofactores que es necesario calcular.

Ejemplo 4.11

Calcule el determinante de

$$A = \begin{bmatrix} 2 & -3 & 0 & 1 \\ 5 & 4 & 2 & 0 \\ 1 & -1 & 0 & 3 \\ -2 & 1 & 0 & 0 \end{bmatrix}$$

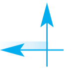
Solución Primero, note que la columna 3 sólo tiene una entrada distinta de cero; por tanto, debe expandir a lo largo de esta columna. A continuación, note que el patrón $+/-$ asigna un signo menos a la entrada $a_{23} = 2$. Por tanto, se tiene

$$\begin{aligned}\det A &= a_{13}C_{13} + a_{23}C_{23} + a_{33}C_{33} + a_{43}C_{43} \\ &= 0(C_{13}) + 2C_{23} + 0(C_{33}) + 0(C_{43}) \\ &= -2\begin{vmatrix} 2 & -3 & 1 \\ 1 & -1 & 3 \\ -2 & 1 & 0 \end{vmatrix}\end{aligned}$$

Ahora continúe expandiendo a lo largo del tercer renglón del determinante anterior (la tercera columna también sería una buena elección) para obtener

$$\begin{aligned}\det A &= -2\left(-2\begin{vmatrix} -3 & 1 \\ -1 & 3 \end{vmatrix} - \begin{vmatrix} 2 & 1 \\ 1 & 3 \end{vmatrix}\right) \\ &= -2(-2(-8) - 5) \\ &= -2(11) = -22\end{aligned}$$

(Note que el patrón $+/-$ para el menor de 3×3 no es el de la matriz original, sino el de una matriz de 3×3 en general).



La expansión de Laplace es particularmente útil cuando la matriz es triangular (superior o inferior).

Ejemplo 4.12 Calcule el determinante de

$$A = \begin{bmatrix} 2 & -3 & 1 & 0 & 4 \\ 0 & 3 & 2 & 5 & 7 \\ 0 & 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

Solución Expanda a lo largo de la primera columna para obtener

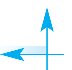
$$\det A = 2\begin{vmatrix} 3 & 2 & 5 & 7 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

(Se omitieron todos los cofactores correspondientes a las entradas cero.) Ahora expanda nuevamente a lo largo de la primera columna:

$$\det A = 2 \cdot 3\begin{vmatrix} 1 & 6 & 0 \\ 0 & 5 & 2 \\ 0 & 0 & -1 \end{vmatrix}$$

Al seguir expandiendo a lo largo de la primera columna, se completa el cálculo:

$$\det A = 2 \cdot 3 \cdot 1\begin{vmatrix} 5 & 2 \\ 0 & -1 \end{vmatrix} = 2 \cdot 3 \cdot 1 \cdot (5(-1) - 2 \cdot 0) = 2 \cdot 3 \cdot 1 \cdot 5 \cdot (-1) = -30$$



El ejemplo 4.12 debe convencerlo de que el determinante de una matriz triangular es el producto de sus entradas diagonales. En el ejercicio 21 se le pedirá ofrecer una prueba de este hecho. El resultado se registra como teorema.

Teorema 4.2 El determinante de una matriz triangular es el producto de las entradas en su diagonal principal. Específicamente, si $A = [a_{ij}]$ es una matriz triangular de $n \times n$, entonces

$$\det A = a_{11}a_{22} \cdots a_{nn}$$

Nota En general (esto es, a menos que la matriz sea triangular o tenga alguna otra forma especial), calcular un determinante mediante expansión por cofactores no es eficiente. Por ejemplo, el determinante de una matriz de 3×3 tiene $6 = 3!$ sumandos, y cada uno requiere dos multiplicaciones, y luego se necesitan cinco sumas y restas para terminar los cálculos. Para una matriz de $n \times n$, habrá $n!$ sumandos, cada uno con $n - 1$ multiplicaciones, y luego $n! - 1$ sumas y restas. El número total de operaciones es por tanto

$$T(n) = (n - 1)n! + n! - 1 > n!$$

Incluso la más rápida de las supercomputadoras no puede calcular el determinante de una matriz moderadamente grande con el uso de la expansión por cofactores. Para ilustrar: suponga que necesita calcular un determinante de 50×50. (Matrices *mucho* más grandes que 50×50 se usan para almacenar los datos de las imágenes digitales como las que se transmiten por Internet o las que se toman con una cámara digital.) Calcular el determinante directamente requeriría, en general, más de 50! operaciones, y $50! \approx 3 \times 10^{64}$. Si tuviera una computadora que pudiera realizar un billón (10^{12}) operaciones por segundo, tardaría aproximadamente 3×10^{52} segundos, o casi 10^{45} años, en terminar los cálculos. Para poner esto en perspectiva, considere que los astrónomos estiman la edad del universo en al menos 10 mil millones (10^{10}) de años. Por ende, incluso en una supercomputadora muy rápida, ¡calcular un determinante 50×50 mediante expansión por cofactores tardaría más de 10^{30} veces la edad del universo!

Por fortuna, existen métodos mejores, y ahora se desarrollarán medios computacionalmente más efectivos para encontrar determinantes. Primero, es necesario observar algunas de las propiedades de los determinantes.

Propiedades de los determinantes

La forma más eficiente de calcular determinantes es usar reducción por renglones. Sin embargo, no toda operación elemental con renglones deja sin cambios el determinante de una matriz. El siguiente teorema resume las principales propiedades que necesita entender con la finalidad de usar de manera efectiva la reducción por renglones.

Teorema 4.3 Sea $A = [a_{ij}]$ una matriz cuadrada.

a. Si A tiene un renglón (columna) cero, entonces $\det A = 0$.
b. Si B se obtiene al intercambiar dos renglones (columnas) de A, entonces $\det B = -\det A$.
c. Si A tiene dos renglones (columnas) idénticos, entonces $\det A = 0$.
d. Si B se obtiene al multiplicar un renglón (columna) de A por k, entonces $\det B = k \det A$.
e. Si A, B y C son idénticas, excepto que el i-ésimo renglón (columna) de C es la suma de los i-ésimos renglones (columnas) de A y B, entonces $\det C = \det A + \det B$.
f. Si B se obtiene al sumar un múltiplo de un renglón (columna) de A a otro renglón (columna), entonces $\det B = \det A$.

Demostración Se demostrará (b) como Lema 4.14 al final de esta sección. Las demostraciones de las propiedades (a) y (f) se dejan como ejercicios. Las propiedades restantes se probarán en términos de renglones; las demostraciones correspondientes a columnas son análogas.

(c) Si A tiene dos renglones idénticos, intercámbielos para obtener la matriz B. Claramente, $B = A$, de modo que $\det B = \det A$. Por otra parte, por (b), $\det B = -\det A$. Por tanto, $\det A = -\det A$, de modo que $\det A = 0$.

(d) Suponga que el renglón i de A se multiplica por k para producir B; esto es, $b_{ij} = ka_{ij}$ para $j = 1, \ldots, n$. Dado que los cofactores C_{ij} de los elementos en los i-ésimos renglones de A y B son idénticos (¿por qué?), expandir a lo largo del i-ésimo renglón de B produce

$$\det B = \sum_{j=1}^{n} b_{ij}C_{ij} = \sum_{j=1}^{n} ka_{ij}C_{ij} = k\sum_{j=1}^{n} a_{ij}C_{ij} = k \det A$$

(e) Como en (d), los cofactores C_{ij} de los elementos en los i-ésimos renglones de A, B y C son idénticos. Más aún, $c_{ij} = a_{ij} + b_{ij}$ para $j = 1, \ldots, n$. Expanda a lo largo del i-ésimo renglón de C para obtener

$$\det C = \sum_{j=1}^{n} c_{ij}C_{ij} = \sum_{j=1}^{n} (a_{ij} + b_{ij})C_{ij} = \sum_{j=1}^{n} a_{ij}C_{ij} + \sum_{j=1}^{n} b_{ij}C_{ij} = \det A + \det B$$

Note que las propiedades (b), (d) y (f) se relacionan con operaciones elementales con renglones. Dado que la forma escalonada de una matriz cuadrada es necesariamente triangular superior, puede combinar estas propiedades con el Teorema 2 para calcular determinantes de manera eficiente. (Vea la Exploración: "Conteo de operaciones" en el capítulo 2, que muestra que la reducción por renglones de una matriz de $n \times n$ usa en el orden de n^3 operaciones, mucho menos que las $n!$ necesarias para la expansión por cofactores.) Los siguientes ejemplos ilustran el cálculo de determinantes usando reducción por renglones.

Ejemplo 4.13

Calcule $\det A$ si

(a) $A = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 5 & 3 \\ -4 & -6 & 2 \end{bmatrix}$

(b) $A = \begin{bmatrix} 0 & 2 & -4 & 5 \\ 3 & 0 & -3 & 6 \\ 2 & 4 & 5 & 7 \\ 5 & -1 & -3 & 1 \end{bmatrix}$

Solución

(a) Con la propiedad (f) y luego la propiedad (a), se tiene

$$\det A = \begin{vmatrix} 2 & 3 & -1 \\ 0 & 5 & 3 \\ -4 & -6 & 2 \end{vmatrix} \overset{R_3+2R_1}{=} \begin{vmatrix} 2 & 3 & -1 \\ 0 & 5 & 3 \\ 0 & 0 & 0 \end{vmatrix} = 0$$

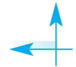
(b) Se reduce A a forma escalonada del modo siguiente (existen otras formas posibles de hacer esto):

$$\det A = \begin{vmatrix} 0 & 2 & -4 & 5 \\ 3 & 0 & -3 & 6 \\ 2 & 4 & 5 & 7 \\ 5 & -1 & -3 & 1 \end{vmatrix} \overset{R_1 \leftrightarrow R_2}{=} - \begin{vmatrix} 3 & 0 & -3 & 6 \\ 0 & 2 & -4 & 5 \\ 2 & 4 & 5 & 7 \\ 5 & -1 & -3 & 1 \end{vmatrix} \overset{R_1/3}{=} -3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & -4 & 5 \\ 2 & 4 & 5 & 7 \\ 5 & -1 & -3 & 1 \end{vmatrix}$$

$$\overset{\substack{R_3-2R_1 \\ R_4-5R_1}}{=} -3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & -4 & 5 \\ 0 & 4 & 7 & 3 \\ 0 & -1 & 2 & -9 \end{vmatrix} \overset{R_2 \leftrightarrow R_4}{=} -(-3) \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & -9 \\ 0 & 4 & 7 & 3 \\ 0 & 2 & -4 & 5 \end{vmatrix}$$

$$\overset{\substack{R_3+4R_2 \\ R_4+2R_2}}{=} 3 \begin{vmatrix} 1 & 0 & -1 & 2 \\ 0 & -1 & 2 & -9 \\ 0 & 0 & 15 & -33 \\ 0 & 0 & 0 & -13 \end{vmatrix}$$

$$= 3 \cdot 1 \cdot (-1) \cdot 15 \cdot (-13) = 585$$



Comentario Por el Teorema 4.3, también puede usar operaciones elementales con columnas en el proceso de calcular determinantes, y puede "mezclar y combinar" operaciones elementales en renglones y columnas. Por ejemplo, en el ejemplo 4.13(a), podría comenzar con la suma de la columna 3 a la columna 1 para crear un 1 como pivote en la esquina superior izquierda. De hecho, el método que se utilizó fue más rápido, pero en otros ejemplos las operaciones en columnas pueden acelerar los cálculos. Tenga esto en mente cuando trabaje determinantes a mano.

Determinantes de matrices elementales

Recuerde de la sección 3.3 que una matriz elemental resulta de realizar una operación elemental con renglón sobre una matriz identidad. Al hacer $A = I_n$ en el Teorema 4.3 se produce el siguiente teorema.

Teorema 4.4 Sea E una matriz elemental de $n \times n$.

a. Si E resulta de intercambiar dos renglones de I_n, entonces $\det E = -1$.
b. Si E resulta de multiplicar un renglón de I_n por k, entonces $\det E = k$.
c. Si E resulta de sumar un múltiplo de un renglón de I_n a otro renglón, entonces $\det E = 1$.

La palabra *lema* se deriva del verbo griego *lambanein*, que significa "asir". En matemáticas, un lema es un "teorema auxiliar" que se "agarra" y se usa para probar otro teorema, por lo general más importante.

Demostración Dado que $\det I_n = 1$, al aplicar (b), (d) y (f) del Teorema 4.3 inmediatamente produce (a), (b) y (c), respectivamente, del Teorema 4.4.

A continuación, recuerde que al multiplicar *por la izquierda* una matriz B por una matriz elemental realiza la correspondiente operación elemental con renglón sobre B. Por tanto, es posible parafrasear sucintamente (b), (d) y (f) del Teorema 4.3 como el lema siguiente, cuya prueba es directa y se deja como el ejercicio 43.

Lema 4.5 Sea B una matriz de $n \times n$ y sea E una matriz elemental de $n \times n$. Entonces

$$\det(EB) = (\det E)(\det B)$$

Puede usar el Lema 4.5 para demostrar el teorema principal de esta sección: una caracterización de la inversibilidad en términos de determinantes.

Teorema 4.6 Una matriz cuadrada A es invertible si y sólo si $\det A \neq 0$.

Demostración Sea A una matriz de $n \times n$ y sea R la forma escalonada reducida por renglones de A. Primero se demostrará que $\det A \neq 0$ si y sólo si $\det R \neq 0$. Sean $E_1, E_2, \ldots, E_r$ las matrices elementales correspondientes a las operaciones elementales con renglones que reducen A a R. Entonces

$$E_r \cdots E_2 E_1 A = R$$

Al sacar determinantes de ambos lados y aplicar repetidamente el Lema 4.5, se obtiene

$$(\det E_r) \cdots (\det E_2)(\det E_1)(\det A) = \det R$$

Por el Teorema 4.4, los determinantes de todas las matrices elementales son distintos de cero. Se concluye que $A \neq 0$ si y sólo si $\det R \neq 0$.

Ahora suponga que A es invertible. Entonces, por el teorema fundamental de las matrices invertibles, $R = I_n$, de modo que $\det R = 1 \neq 0$. En consecuencia, también $\det A \neq 0$. Por el contrario, si $A \neq 0$, entonces $\det R \neq 0$, de modo que R no puede contener un renglón cero, por el Teorema 4.3(a). Se tiene que R debe ser I_n (¿por qué?), de modo que A es invertible, nuevamente por el teorema fundamental.

Determinantes y operaciones matriciales

Ahora se tratará de determinar qué relación, si la hay, existe entre los determinantes y algunas de las operaciones matriciales básicas. Específicamente, se quieren encontrar fórmulas para $\det(kA)$, $\det(A + B)$, $\det(AB)$ $\det(A^{-1})$, y $\det(A^T)$ en términos de $\det A$ y $\det B$.

El Teorema 4.3(d) *no* dice que $\det(kA) = k \det A$. La relación correcta entre la multiplicación escalar y los determinantes está dada por el siguiente teorema.

Teorema 4.7 Si A es una matriz de $n \times n$, entonces

$$\det(kA) = k^n \det A$$

En el ejercicio 44 se le pide dar una demostración de este teorema.

Por desgracia, no hay una fórmula simple para $\det(A + B)$ y, en general, $\det(A + B) \neq \det A + \det B$. (Encuentre dos matrices de 2×2 que verifiquen esto.) En consecuencia, es una agradable sorpresa descubrir que los determinantes son bastante compatibles con la multiplicación de matrices. De hecho, se tiene la agradable fórmula siguiente debida a Cauchy.

Augustin Louis Cauchy (1789–1857) nació en París y estudió ingeniería, pero cambió a matemáticas debido a su pobre salud. Brillante y prolífico matemático, publicó más de 700 artículos, muchos acerca de problemas bastante difíciles. Su nombre puede encontrarse en muchos teoremas y definiciones en ecuaciones diferenciales, series infinitas, teoría de probabilidad, álgebra y física. Es célebre por introducir el rigor en el cálculo y tender los cimientos de la rama de las matemáticas conocida como análisis. Políticamente conservador, Cauchy fue realista, y en 1830 siguió a Carlos X al exilio. Regresó a Francia en 1838, pero no recuperó su puesto en la Sorbona hasta que la universidad desechó su requisito de que los miembros del personal docente juraran lealtad al nuevo rey.

Teorema 4.8 Si A y B son matrices de $n \times n$, entonces

$$\det(AB) = (\det A)(\det B)$$

Demostración Considere dos casos: A invertible y A no invertible.

Si A es invertible, entonces, por el teorema fundamental de las matrices invertibles, puede escribirse como un producto de matrices elementales, por ejemplo

$$A = E_1E_2 \cdots E_k$$

Entonces $AB = E_1E_2 \cdots E_kB$, de modo que k aplicaciones del Lema 4.5 producirán

$$\det(AB) = \det(E_1E_2 \cdots E_kB) = (\det E_1)(\det E_2) \cdots (\det E_k)(\det B)$$

Al continuar con la aplicación del Lema 4.5, se obtiene

$$\det(AB) = \det(E_1E_2 \cdots E_k)\det B = (\det A)(\det B)$$

Por otra parte, si A no es invertible, entonces tampoco lo es AB, por el ejercicio 47 de la sección 3.3. En consecuencia, por el Teorema 4.6, $\det A = 0$ y $\det(AB) = 0$. En consecuencia, $\det(AB) = (\det A)(\det B)$, pues ambos lados son cero.

Ejemplo 4.14

Al aplicar el Teorema 4.8 a $A = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}$ y $B = \begin{bmatrix} 5 & 1 \\ 2 & 1 \end{bmatrix}$, se encuentra que

$$AB = \begin{bmatrix} 12 & 3 \\ 16 & 5 \end{bmatrix}$$

y que $\det A = 4$, $\det B = 3$ y $\det(AB) = 12 = 4 \cdot 3 = (\det A)(\det B)$, como se afirmó. (¡Compruebe estas aseveraciones!)

El siguiente teorema ofrece una amigable relación entre el determinante de una matriz invertible y el determinante de su inversa.

Teorema 4.9 Si A es invertible, entonces

$$\det(A^{-1}) = \frac{1}{\det A}$$

Demostración Puesto que A es invertible, $AA^{-1} = I$, de modo que $\det(AA^{-1}) = \det I = 1$. En consecuencia, $(\det A)(\det A^{-1}) = 1$, por el Teorema 4.8, y como $\det A \neq 0$ (¿por qué?), dividir entre $\det A$ produce el resultado.

Ejemplo 4.15 Verifique el Teorema 4.9 para la matriz A del ejemplo 4.14.

Solución Calcule

$$A^{-1} = \tfrac{1}{4}\begin{bmatrix} 3 & -1 \\ -2 & 2 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} & -\frac{1}{4} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

de modo que

$$\det A^{-1} = \left(\frac{3}{4}\right)\left(\frac{1}{2}\right) - \left(-\frac{1}{4}\right)\left(-\frac{1}{2}\right) = \frac{3}{8} - \frac{1}{8} = \frac{1}{4} = \frac{1}{\det A}$$

Comentario La belleza del Teorema 4.9 es que en ocasiones no es necesario saber *cuál* es la inversa de una matriz, sino sólo que *existe*, o saber cuál es su determinante. Para la matriz A en los dos últimos ejemplos, una vez que se sabe que $\det A = 4 \neq 0$, inmediatamente puede deducirse que A es invertible y que $\det A^{-1} = \frac{1}{4}$ sin realmente calcular A^{-1}.

Ahora se relacionará el determinante de una matriz A con el de su transpuesta A^T. Dado que los renglones de A^T son justo las columnas de A, evaluar $\det A^T$ mediante expansión a lo largo del primer renglón es idéntico a evaluar $\det A$ por expansión a lo largo de su primera columna, lo cual es posible hacer por el teorema de expansión de Laplace. Por tanto, se tiene el siguiente resultado.

Teorema 4.10 Para cualquier matriz cuadrada A,

$$\det A = \det A^T$$

Regla de Cramer y la adjunta

Gabriel Cramer (1704-1752) fue un matemático suizo. La regla que lleva su nombre se publicó en 1750, en su tratado *Introducción al análisis de las curvas algebraicas*. Sin embargo, ya en 1730, otros matemáticos conocían casos especiales de la fórmula, entre ellos el escocés Colin Maclaurin (1698-1746), acaso el más grande de los matemáticos británicos que fueron los "sucesores de Newton".

En esta sección se deducen dos fórmulas útiles que relacionan los determinantes con la solución de sistemas lineales y la inversa de una matriz. La primera de ellas, la ***regla de Cramer***, ofrece una fórmula para describir la solución de ciertos sistemas de n ecuaciones lineales con n variables completamente en términos de determinantes. Aunque este resultado es de poco uso práctico más allá de sistemas de 2×2, es de gran importancia teórica.

Se necesitará una nueva notación para este resultado y su prueba. Para una matriz A de $n \times n$ y un vector $\mathbf{b}$ en $\mathbb{R}^n$, sea $A_i(\mathbf{b})$ que denota la matriz obtenida al sustituir la i-ésima columna de A por $\mathbf{b}$. Esto es:

$$A_i(\mathbf{b}) = [\mathbf{a}_1 \cdots \underset{\substack{\uparrow \\ \text{Columna } i}}{\mathbf{b}} \cdots \mathbf{a}_n]$$

Teorema 4.11 **Regla de Cramer**

Sea A una matriz de $n \times n$ invertible y sea $\mathbf{b}$ un vector en $\mathbb{R}^n$. Entonces la solución única $\mathbf{x}$ del sistema $A\mathbf{x} = \mathbf{b}$ está dada por

$$x_i = \frac{\det(A_i(\mathbf{b}))}{\det A} \quad \text{para } i = 1, \ldots, n$$

Demostración Las columnas de la matriz identidad $I = I_n$ son los vectores unitarios estándar $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Si $A\mathbf{x} = \mathbf{b}$, entonces

$$\begin{aligned} AI_i(\mathbf{x}) &= A[\mathbf{e}_1 \ \cdots \ \mathbf{x} \ \cdots \ \mathbf{e}_n] = [A\mathbf{e}_1 \ \cdots \ A\mathbf{x} \ \cdots \ A\mathbf{e}_n] \\ &= [\mathbf{a}_1 \ \cdots \ \mathbf{b} \ \cdots \ \mathbf{a}_n] = A_i(\mathbf{b}) \end{aligned}$$

Por tanto, por el Teorema 4.8,

$$(\det A)(\det I_i(\mathbf{x})) = \det(AI_i(\mathbf{x})) = \det(A_i(\mathbf{b}))$$

Ahora

$$\det I_i(\mathbf{x}) = \begin{vmatrix} 1 & 0 & \cdots & x_1 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & x_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & x_i & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & x_{n-1} & \cdots & 1 & 0 \\ 0 & 0 & \cdots & x_n & \cdots & 0 & 1 \end{vmatrix} = x_i$$

como puede ver al expandir a lo largo del i-ésimo renglón. Por tanto, $(\det A)x_i = \det(A_i(\mathbf{b}))$ y el resultado surge al dividir entre $\det A$ (que es distinto de cero, pues A es invertible).

Ejemplo 4.16 Use la Regla de Cramer para resolver el sistema

$$\begin{aligned} x_1 + 2x_2 &= 2 \\ -x_1 + 4x_2 &= 1 \end{aligned}$$

Solución Calcule

$$\det A = \begin{vmatrix} 1 & 2 \\ -1 & 4 \end{vmatrix} = 6, \quad \det(A_1(\mathbf{b})) = \begin{vmatrix} 2 & 2 \\ 1 & 4 \end{vmatrix} = 6 \quad \text{y} \quad \det(A_2(\mathbf{b})) = \begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix} = 3$$

Por la Regla de Cramer,

$$x_1 = \frac{\det(A_1(\mathbf{b}))}{\det A} = \frac{6}{6} = 1 \quad \text{y} \quad x_2 = \frac{\det(A_2(\mathbf{b}))}{\det A} = \frac{3}{6} = \frac{1}{2}$$

Comentario Como se hizo notar anteriormente, la regla de Cramer es computacionalmente ineficiente para todos los sistemas de ecuaciones lineales, excepto los pequeños, pues involucra el cálculo de muchos determinantes. El esfuerzo que se realiza para calcular sólo uno de dichos determinantes, usando incluso el método más eficiente, se emplearía mejor al usar la eliminación gaussiana para resolver directamente el sistema.

El resultado final de esta sección es una fórmula para la inversa de una matriz en términos de determinantes. Esta fórmula se insinuó por la fórmula para la inversa de una matriz de 3×3, que se dio sin demostración al inicio de esta sección. Por ende, se cierra el ciclo.

Ahora descubra la fórmula por usted mismo. Si A es una matriz invertible de $n \times n$, su inversa es la matriz (única) X que satisface la ecuación $AX = I$. Al resolver para X una columna cada vez, sea $\mathbf{x}_j$ la j-ésima columna de X. Esto es:

$$\mathbf{x}_j = \begin{bmatrix} x_{1j} \\ \vdots \\ x_{ij} \\ \vdots \\ x_{nj} \end{bmatrix}$$

Por tanto, $A\mathbf{x}_j = \mathbf{e}_j$, y por la regla de Cramer,

$$x_{ij} = \frac{\det(A_i(\mathbf{e}_j))}{\det A}$$

Sin embargo,

$$\det(A_i(\mathbf{e}_j)) = \begin{vmatrix} a_{11} & a_{12} & \cdots & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \iddots & \vdots \\ a_{j1} & a_{j2} & \cdots & 1 & \cdots & a_{jn} \\ \vdots & \vdots & \iddots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 & \cdots & a_{nn} \end{vmatrix} = (-1)^{j+i} \det A_{ji} = C_{ji}$$

(i-ésima columna: la columna que contiene $0, 0, \ldots, 1, \ldots, 0$)

que es el cofactor (j, i) de A.

Se tiene que $x_{ij} = (1/\det A)C_{ji}$, de modo que $A^{-1} = X = (1/\det A)[C_{ji}] = (1/\det A)[C_{ij}]^T$. Con palabras: la inversa de A es la *transpuesta* de la matriz de cofactores de A, divida por el determinante de A.

La matriz

$$[C_{ji}] = [C_{ij}]^T = \begin{bmatrix} C_{11} & C_{21} & \cdots & C_{n1} \\ C_{12} & C_{22} & \cdots & C_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ C_{1n} & C_{2n} & \cdots & C_{nn} \end{bmatrix}$$

se llama ***adjunta*** (o ***transpuesta conjugada***) de A y se denota adj A. El resultado que acaba de probarse puede enunciarse del modo siguiente.

Teorema 4.12 Sea A una matriz de $n \times n$ invertible. Entonces

$$A^{-1} = \frac{1}{\det A}\operatorname{adj} A$$

Ejemplo 4.17 Use el método de la adjunta para calcular la inversa de

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 2 & 4 \\ 1 & 3 & -3 \end{bmatrix}$$

Solución Calcule $\det A = -2$ y los nueve cofactores

$$C_{11} = +\begin{vmatrix} 2 & 4 \\ 3 & -3 \end{vmatrix} = -18 \quad C_{12} = -\begin{vmatrix} 2 & 4 \\ 1 & -3 \end{vmatrix} = 10 \quad C_{13} = +\begin{vmatrix} 2 & 2 \\ 1 & 3 \end{vmatrix} = 4$$

$$C_{21} = -\begin{vmatrix} 2 & -1 \\ 3 & -3 \end{vmatrix} = 3 \quad C_{22} = +\begin{vmatrix} 1 & -1 \\ 1 & -3 \end{vmatrix} = -2 \quad C_{23} = -\begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix} = -1$$

$$C_{31} = +\begin{vmatrix} 2 & -1 \\ 2 & 4 \end{vmatrix} = 10 \quad C_{32} = -\begin{vmatrix} 1 & -1 \\ 2 & 4 \end{vmatrix} = -6 \quad C_{33} = +\begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix} = -2$$

La adjunta es la *transpuesta* de la matriz de cofactores; a saber,

$$\operatorname{adj} A = \begin{bmatrix} -18 & 10 & 4 \\ 3 & -2 & -1 \\ 10 & -6 & -2 \end{bmatrix}^T = \begin{bmatrix} -18 & 3 & 10 \\ 10 & -2 & -6 \\ 4 & -1 & -2 \end{bmatrix}$$

Entonces

$$A^{-1} = \frac{1}{\det A}\operatorname{adj} A = -\frac{1}{2}\begin{bmatrix} -18 & 3 & 10 \\ 10 & -2 & -6 \\ 4 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 9 & -\frac{3}{2} & -5 \\ -5 & 1 & 3 \\ -2 & \frac{1}{2} & 1 \end{bmatrix}$$

que es la misma respuesta que se obtuvo (con menos trabajo) en el ejemplo 3.30.

Demostración del teorema de expansión de Laplace

Por desgracia no hay una demostración sencilla y breve del Teorema de Expansión de Laplace. La demostración que se proporciona tiene el mérito de ser relativamente directa. Se le descompone en varios pasos, el primero de los cuales es demostrar que la expansión por cofactores a lo largo del primer renglón de una matriz es igual que la expansión por cofactores a lo largo de la primera columna.

Lema 4.13 Sea A una matriz de $n \times n$. Entonces

$$a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} = \det A = a_{11}C_{11} + a_{21}C_{21} + \cdots + a_{n1}C_{n1} \qquad (7)$$

Demostración Este lema se demuestra por inducción sobre n. Para $n = 1$, el resultado es trivial. Ahora suponga que el resultado es verdadero para matrices $(n - 1) \times (n - 1)$; esta es la hipótesis de inducción. Note que, por la definición de cofactor (o menor), todos los términos que contienen a_{11} se representan por medio del sumando $a_{11}C_{11}$. Por tanto, puede ignorar los términos que contengan a_{11}.

El i-ésimo sumando a la derecha de la ecuación (7) es $a_{i1}C_{i1} = a_{i1}(-1)^{i+1} \det A_{i1}$. Ahora se expande A_{i1} a lo largo del primer renglón:

$$\begin{vmatrix} a_{12} & a_{13} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots & \iddots & \vdots \\ a_{i-1,2} & a_{i-1,3} & \cdots & a_{i-1,j} & \cdots & a_{i-1,n} \\ a_{i+1,2} & a_{i+1,3} & \cdots & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \iddots & \vdots & \ddots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{nj} & \cdots & a_{n,n} \end{vmatrix}$$

El j-ésimo término de esta expansión de $\det A_{i1}$ es $a_{1j}(-1)^{1+j-1} \det A_{1i,1j}$, donde la notación $A_{kl,rs}$ denota la submatriz de A que se obtiene al borrar los renglones k y l y las columnas r y s. Al combinar esto, se ve que el término que contiene $a_{i1}a_{1j}$ en el lado derecho de la ecuación (7) es

$$a_{i1}(-1)^{i+1}a_{1j}(-1)^{1+j-1} \det A_{1i,1j} = (-1)^{i+j+1}a_{i1}a_{1j} \det A_{1i,1j}$$

¿Cuál es el término que contiene $a_{i1}a_{1j}$ en el lado izquierdo de la ecuación (7)? El factor a_{1j} ocurre en el j-ésimo sumando $a_{1j}C_{1j} = a_{1j}(-1)^{1+j} \det A_{1j}$. Por la hipótesis de inducción, puede expandir $\det A_{ij}$ a lo largo de su primera columna:

$$\begin{vmatrix} a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ a_{31} & \cdots & a_{3,j-1} & a_{3,j+1} & \cdots & a_{3n} \\ \vdots & \ddots & \vdots & \vdots & \iddots & \vdots \\ a_{i1} & \cdots & a_{i,j-1} & a_{i,j+1} & \cdots & a_{in} \\ \vdots & \iddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

El i-ésimo término en esta expansión de $\det A_{1j}$ es $a_{i1}(-1)^{(i-1)+1} \det A_{1i,1j}$ de modo que el término que contiene $a_{i1}a_{1j}$ en el lado izquierdo de la ecuación (7) es

$$a_{1j}(-1)^{1+j}a_{i1}(-1)^{(i-1)+1} \det A_{1i,1j} = (-1)^{i+j+1}a_{i1}a_{1j} \det A_{1i,1j}$$

que establece que los lados izquierdo y derecho de la ecuación (7) son equivalentes.

A continuación se demuestra la propiedad (b) del Teorema 4.3.

Lema 4.14 Sea A una matriz de $n \times n$ y sea B que se obtiene al intercambiar cualesquiera dos renglones (columnas) de A. Entonces

$$\det B = -\det A$$

Demostración Una vez más, la demostración es por inducción sobre n. El resultado puede comprobarse fácilmente cuando $n = 2$, así que suponga que es verdadero para matrices $(n-1)\times(n-1)$. Se demostrará que el resultado es verdadero para matrices de $n \times n$. Primero, pruebe que se cumple cuando se intercambian dos renglones adyacentes de A; por ejemplo, los renglones r y $r+1$.

Por el Lema 4.13, puede evaluar $\det B$ mediante expansión por cofactores a lo largo de su primera columna. El i-ésimo término en esta expansión es $(-1)^{1+i}b_{i1}\det B_{i1}$. Si $i \neq r$ e $i \neq r+1$, entonces $b_{i1} = a_{i1}$ y B_{i1} es una submatriz $(n-1)\times(n-1)$ que es idéntica a A_{i1}, excepto que se intercambiaron dos renglones adyacentes.

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{r+1,1} & a_{r+1,2} & \cdots & a_{r+1,n} \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Por tanto, por la hipótesis de inducción, $\det B_{i1} = -\det A_{i1}$ si $i \neq r$ e $i \neq r+1$.

Si $i = r$, entonces $b_{i1} = a_{r+1,1}$ y $B_{i1} = A_{r+1,1}$.

$$\text{Renglón } i \rightarrow \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{r+1,1} & a_{r+1,2} & \cdots & a_{r+1,n} \\ a_{r1} & a_{r2} & \cdots & a_{rn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

Por tanto, el r-ésimo sumando en $\det B$ es

$$(-1)^{r+1}b_{r1}\det B_{r1} = (-1)^{r+1}a_{r+1,1}\det A_{r+1,1} = -(-1)^{(r+1)+1}a_{r+1,1}\det A_{r+1,1}$$

De igual modo, si $i = r+1$, entonces $b_{i1} = a_{r1}$, $B_{i1} = A_{r1}$ y el $(r+1)$-ésimo sumando en el $\det B$ es

$$(-1)^{(r+1)+1}b_{r+1,1}\det B_{r+1,1} = (-1)^{r}a_{r1}\det A_{r1} = -(-1)^{r+1}a_{r1}\det A_{r1}$$

En otras palabras, los términos r-ésimo y $(r+1)$-ésimo en la expansión por cofactores de la primera columna de $\det B$ son los *negativos* de los términos $(r+1)$-ésimo y r-ésimo, respectivamente, en la expansión por cofactores de la primera columna de $\det A$.

Al sustituir todos estos resultados en el $\det B$ y usar el Lema 4.13 nuevamente, se obtiene

$$\begin{aligned} \det B &= \sum_{i=1}^{n}(-1)^{i+1}b_{i1}\det B_{i1} \\ &= \sum_{\substack{i=1 \\ i\neq r,r+1}}^{n}(-1)^{i+1}b_{i1}\det B_{i1} + (-1)^{r+1}b_{r1}\det B_{r1} + (-1)^{(r+1)+1}b_{r+1,1}\det B_{r+1,1} \\ &= \sum_{\substack{i=1 \\ i\neq r,r+1}}^{n}(-1)^{i+1}a_{i1}(-\det A_{i1}) - (-1)^{(r+1)+1}a_{r+1,1}\det A_{r+1,1} - (-1)^{r+1}a_{r1}\det A_{r1} \\ &= -\sum_{i=1}^{n}(-1)^{i+1}a_{i1}\det A_{i1} \\ &= -\det A \end{aligned}$$

Esto demuestra el resultado para matrices de $n \times n$ si se intercambian renglones adyacentes. Para ver que se sustenta para intercambios arbitrarios de renglones, sólo es necesario notar que, por ejemplo, los renglones r y s, donde $r < s$, pueden intercambiarse al realizar $2(s - r) - 1$ intercambios de renglones adyacentes (vea el ejercicio 67). Puesto que el número de intercambios es *impar* y cada uno cambia el signo del determinante, el efecto neto es un cambio de signo, como se desea.

La demostración para los intercambios de columna es análoga, excepto que se expande a lo largo del renglón 1 en lugar de a lo largo de la columna 1.

Ahora se puede demostrar el teorema de expansión de Laplace.

Demostración del Teorema 4.1 Sea B la matriz que se obtiene al mover el renglón i de A a la parte superior, usando $i - 1$ intercambios de renglones adyacentes. Por el Lema 4.14, $\det B = (-1)^{i-1} \det A$. Pero $b_{1j} = a_{ij}$ y $B_{1j} = A_{ij}$ para $j = 1, \ldots, n$.

$$\det B = \begin{vmatrix} a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

Por tanto,

$$\det A = (-1)^{i-1} \det B = (-1)^{i-1} \sum_{j=1}^{n} (-1)^{1+j} b_{1j} \det B_{1j}$$

$$= (-1)^{i-1} \sum_{j=1}^{n} (-1)^{1+j} a_{ij} \det A_{ij} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$$

lo que da la fórmula para la expansión por cofactores a lo largo del renglón i.

La demostración para expansión por columna es similar e invoca el Lema 4.13, de modo que es posible usar expansión por columna en lugar de expansión por renglón (vea el ejercicio 68).

Breve historia de los determinantes

Como se anotó al inicio de esta sección, la historia de los determinantes antecede a la de las matrices. De hecho, los determinantes se introdujeron primero, de manera independiente, por parte de Seki, en 1683, y Leibniz, en 1693. En 1748, los determinantes aparecieron en el *Tratado de álgebra* de Maclaurin, que incluyó un tratamiento de la regla de Cramer hasta el caso de 4×4. En 1750, el mismo Cramer probó el caso general de su regla y la aplicó al ajuste de curvas, y en 1772 Laplace dio una demostración de su teorema de expansión.

Un niño prodigio autodidacta, Takakazu Seki Kōwa (1642-1708), descendía de una familia de guerreros samurai. Además de descubrir los determinantes, escribió acerca de ecuaciones diofantinas, cuadrados mágicos y números de Bernoulli (antes que Bernoulli) y muy probablemente realizó descubrimientos en cálculo.

El término *determinante* no se acuñó sino hasta 1801, cuando lo usó Gauss. En 1812, Cauchy usó por primera vez los determinantes en el sentido moderno. De hecho, Cauchy fue el responsable de desarrollar mucha de la teoría temprana de los determinantes, incluidos muchos resultados importantes que se han mencionado: la regla del producto para determinantes, el polinomio característico y la noción de una matriz diagonalizable. Los determinantes no fueron ampliamente conocidos sino hasta 1841, cuando Jacobi los popularizó, aunque en el contexto de funciones de varias variables, como se les encuentra en un curso de cálculo multivariable. (En 1850, Sylvester llamó “jacobianos” a ese tipo de determinantes, un término que todavía se usa en la actualidad.)

Gottfried Wilhelm von Leibniz (1646-1716) nació en Leipzig y estudió leyes, teología, filosofía y matemáticas. Probablemente es mejor conocido por desarrollar (con Newton, de manera independiente) las ideas principales del cálculo diferencial e integral. Sin embargo, sus aportaciones a otras ramas de las matemáticas también son impresionantes. Desarrolló la noción de determinante, conoció versiones de la regla de Cramer y del teorema de expansión de Laplace antes de que otros le dieran el crédito a ellos, y tendió los cimientos para la teoría de matrices mediante el trabajo que hizo sobre las formas cuadráticas. Leibniz también fue el primero en desarrollar el sistema binario de aritmética. Creía en la importancia de una buena notación y, junto con la notación familiar para derivadas e integrales, introdujo una forma de notación de subíndices para los coeficientes de un sistema lineal que en esencia es la que se usa en la actualidad.

Hacia finales del siglo XIX, la teoría de los determinantes se desarrolló hasta el punto de que libros completos se dedicaban a ellos, incluido el *Tratado elemental sobre los determinantes* de Dodgson, en 1867, y la monumental obra en cinco volúmenes de Thomas Muir, que apareció a principios del siglo XX. Aunque su historia es fascinante, los determinantes de hoy son más de interés teórico que práctico. La Regla de Cramer es un método desesperadamente ineficiente para resolver un sistema de ecuaciones lineales, y los métodos numéricos han sustituido cualquier uso de los determinantes en el cálculo de eigenvalores. Sin embargo, los determinantes se usan para dar a los estudiantes una comprensión inicial de los polinomios característicos (como en las secciones 4.1 y 4.3).

Ejercicios 4.2

Calcule los determinantes en los ejercicios 1-6 usando expansión por cofactores a lo largo del primer renglón y la primera columna.

1. $\begin{vmatrix} 1 & 0 & 3 \\ 5 & 1 & 1 \\ 0 & 1 & 2 \end{vmatrix}$

2. $\begin{vmatrix} 1 & 0 & -2 \\ 3 & 3 & 2 \\ 0 & -1 & 1 \end{vmatrix}$

3. $\begin{vmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{vmatrix}$

4. $\begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{vmatrix}$

5. $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{vmatrix}$

6. $\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{vmatrix}$

Calcule los determinantes en los ejercicios 7-15 usando expansión por cofactores a lo largo de cualquier renglón o columna que parezca conveniente.

7. $\begin{vmatrix} 5 & 2 & 2 \\ -1 & 1 & 2 \\ 3 & 0 & 0 \end{vmatrix}$

8. $\begin{vmatrix} 1 & 2 & 3 \\ -4 & 0 & 4 \\ -3 & -2 & -1 \end{vmatrix}$

9. $\begin{vmatrix} -4 & 1 & 3 \\ 2 & -2 & 4 \\ 1 & -1 & 0 \end{vmatrix}$

10. $\begin{vmatrix} 0 & 0 & \operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta & \tan\theta \\ -\cos\theta & \operatorname{sen}\theta & -\cos\theta \end{vmatrix}$

11. $\begin{vmatrix} a & b & 0 \\ 0 & a & b \\ a & 0 & b \end{vmatrix}$

12. $\begin{vmatrix} 0 & a & 0 \\ b & c & d \\ 0 & e & 0 \end{vmatrix}$

13. $\begin{vmatrix} 1 & -1 & 0 & 3 \\ 2 & 5 & 2 & 6 \\ 0 & 1 & 0 & 0 \\ 1 & 4 & 2 & 1 \end{vmatrix}$

14. $\begin{vmatrix} 3 & -2 & 0 & 1 \\ 1 & 3 & 0 & -1 \\ 0 & 2 & 2 & 4 \\ 3 & 1 & 0 & 0 \end{vmatrix}$

15. $\begin{vmatrix} 0 & 0 & 0 & a \\ 0 & 0 & b & c \\ 0 & d & e & f \\ g & h & i & j \end{vmatrix}$

En los ejercicios 16-18, calcule los determinantes de 3 × 3 indicados, usando el método del ejemplo 4.9.

16. El determinante en el ejercicio 6

17. El determinante en el ejercicio 8

18. El determinante en el ejercicio 11

19. Verifique que el método indicado en (2) concuerda con la ecuación (1) para un determinante de 3 × 3.

20. Verifique que la definición (4) concuerda con la definición de un determinante de 2 × 2 cuando $n = 2$.

21. Demuestre el Teorema 4.2. [*Sugerencia:* una demostración por inducción sería adecuada.]

En los ejercicios 22-25, evalúe el determinante dado usando operaciones elementales en renglón y/o columna y el Teorema 4.3 para reducir la matriz a forma escalonada por renglones.

22. El determinante en el ejercicio 1

23. El determinante en el ejercicio 9

24. El determinante en el ejercicio 13

25. El determinante en el ejercicio 14

En los ejercicios 26-34, use las propiedades de los determinantes para evaluar por inspección el determinante dado. Explique su razonamiento.

26. $\begin{vmatrix} 1 & 1 & 1 \\ 3 & 0 & -2 \\ 2 & 2 & 2 \end{vmatrix}$

27. $\begin{vmatrix} 3 & 1 & 0 \\ 0 & -2 & 5 \\ 0 & 0 & 4 \end{vmatrix}$

28. $\begin{vmatrix} 0 & 0 & 1 \\ 0 & 5 & 2 \\ 3 & -1 & 4 \end{vmatrix}$

29. $\begin{vmatrix} 2 & 3 & -4 \\ 1 & -3 & -2 \\ -1 & 5 & 2 \end{vmatrix}$

30. $\begin{vmatrix} 1 & 2 & 3 \\ 0 & 4 & 1 \\ 1 & 6 & 4 \end{vmatrix}$

31. $\begin{vmatrix} 4 & 1 & 3 \\ -2 & 0 & -2 \\ 5 & 4 & 1 \end{vmatrix}$

32. $\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$

33. $\begin{vmatrix} 0 & 2 & 0 & 0 \\ -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 \end{vmatrix}$

34. $\begin{vmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{vmatrix}$

Encuentre los determinantes en los ejercicios 35-40, suponiendo que

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 4$$

35. $\begin{vmatrix} 2a & 2b & 2c \\ d & e & f \\ g & h & i \end{vmatrix}$

36. $\begin{vmatrix} 3a & -b & 2c \\ 3d & -e & 2f \\ 3g & -h & 2i \end{vmatrix}$

37. $\begin{vmatrix} d & e & f \\ a & b & c \\ g & h & i \end{vmatrix}$

38. $\begin{vmatrix} a+g & b+h & c+i \\ d & e & f \\ g & h & i \end{vmatrix}$

39. $\begin{vmatrix} 2c & b & a \\ 2f & e & d \\ 2i & h & g \end{vmatrix}$

40. $\begin{vmatrix} a & b & c \\ 2d-3g & 2e-3h & 2f-3i \\ g & h & i \end{vmatrix}$

41. Demuestre el Teorema 4.3(a).

42. Demuestre el Teorema 4.3(f).

43. Demuestre el Lema 4.5.

44. Demuestre el Teorema 4.7.

En los ejercicios 45 y 46, use el Teorema 4.6 para encontrar todos los valores de k para los cuales A es invertible.

45. $A = \begin{bmatrix} k & -k & 3 \\ 0 & k+1 & 1 \\ k & -8 & k-1 \end{bmatrix}$

46. $A = \begin{bmatrix} k & k & 0 \\ k^2 & 4 & k^2 \\ 0 & k & k \end{bmatrix}$

En los ejercicios 47-52, suponga que A y B son matrices de n × n con det A = 3 y det B = −2. Encuentre los determinantes indicados.

47. $\det(AB)$ **48.** $\det(A^2)$ **49.** $\det(B^{-1}A)$

50. $\det(2A)$ **51.** $\det(3B^T)$ **52.** $\det(AA^T)$

En los ejercicios 53-56, A y B son matrices de n × n.

53. Demuestre que $\det(AB) = \det(BA)$.

54. Si B es invertible, demuestre que $\det(B^{-1}AB) = \det(A)$.

55. Si A es idempotente (esto es, $A^2 = A$), encuentre todos los posibles valores de $\det(A)$.

56. Una matriz cuadrada A se llama ***nilpotente*** si $A^m = O$ para algún $m > 1$. (La palabra *nilpotente* proviene del latín *nil*, que significa "nada", y *potere*, que significa

"tener poder". En consecuencia, una matriz nilpotente es aquella que se vuelve "nada" [esto es: la matriz cero] cuando se eleva a alguna potencia.) Encuentre todos los posibles valores de $\det(A)$ si A es nilpotente.

En los ejercicios 57-60, use la regla de Cramer para resolver el sistema lineal dado.

57. $\begin{aligned} x + y &= 1 \\ x - y &= 2 \end{aligned}$

58. $\begin{aligned} 2x - y &= 5 \\ x + 3y &= -1 \end{aligned}$

59. $\begin{aligned} 2x + y + 3z &= 1 \\ y + z &= 1 \\ z &= 1 \end{aligned}$

60. $\begin{aligned} x + y - z &= 1 \\ x + y + z &= 2 \\ x - y \quad &= 3 \end{aligned}$

En los ejercicios 61-64, use el Teorema 4.12 para calcular la inversa de la matriz de coeficientes para el ejercicio dado.

61. Ejercicio 57

62. Ejercicio 58

63. Ejercicio 59

64. Ejercicio 60

65. Si A es una matriz invertible de $n \times n$, demuestre que adj A también es invertible y que

$$(\text{adj } A)^{-1} = \frac{1}{\det A} A = \text{adj}\,(A^{-1})$$

66. Si A es una matriz de $n \times n$, demuestre que

$$\det(\text{adj } A) = (\det A)^{n-1}$$

67. Verifique que si $r < s$, entonces los renglones r y s de una matriz pueden intercambiarse al realizar $2(s - r) - 1$ intercambios de renglones adyacentes.

68. Demuestre que el teorema de expansión de Laplace se sostiene para expansión por columna a lo largo de la j-ésima columna.

69. Sea A una matriz cuadrada que puede particionarse como

$$A = \left[\begin{array}{c|c} P & Q \\ \hline O & S \end{array}\right]$$

donde P y S son matrices cuadradas. Se dice que tal matriz está en ***forma triangular (superior) en bloques***. Demuestre que

$$\det A = (\det P)(\det S)$$

[*Sugerencia:* intente una demostración por inducción acerca del número de renglones de P.]

70. **(a)** Ofrezca un ejemplo para demostrar que si A puede particionarse como

$$A = \left[\begin{array}{c|c} P & Q \\ \hline R & S \end{array}\right]$$

donde P, Q, R y S son todas cuadradas, entonces no necesariamente es verdadero que

$$\det A = (\det P)(\det S) - (\det Q)(\det R)$$

(b) Suponga que A se particiona como en el inciso (a) y que P es invertible. Sea

$$B = \left[\begin{array}{c|c} P^{-1} & O \\ \hline -RP^{-1} & I \end{array}\right]$$

Calcule $\det(BA)$ usando el ejercicio 69 y use el resultado para demostrar que

$$\det A = \det P \det(S - RP^{-1}Q)$$

[La matriz $S - RP^{-1}Q$ se llama ***complemento de Schur*** de P en A, en honor a Issai Schur (1875-1941), quien nació en Bielorrusia pero pasó la mayor parte de su vida en Alemania. Es conocido principalmente por su trabajo fundamental acerca de la teoría de la representación de grupos, pero también trabajó en teoría de números, análisis y otras áreas.]

(c) Suponga que A se particiona como en el inciso (a), que P es invertible y que $PR = RP$. Demuestre que

$$\det A = \det(PS - RQ)$$

Método de condensación de Lewis Carroll

Charles Lutwidge Dodgson (1832-1898) es mucho mejor conocido por su nombre literario, Lewis Carroll, con el cual escribió *Las aventuras de Alicia en el país de las maravillas* y *A través del espejo*. También escribió muchos libros matemáticos y colecciones de acertijos lógicos.

En 1866, Charles Dodgson, mejor conocido por su pseudónimo, Lewis Carroll, publicó su único artículo de investigación matemática. En él describió un "nuevo y breve método" para calcular determinantes, al que llamó "condensación". Aunque no es muy conocido en la actualidad y se volvió obsoleto por los métodos numéricos para evaluar determinantes, el método de condensación es muy útil para cálculos manuales. Cuando no había calculadoras o sistemas algebraicos de cómputo, muchos estudiantes descubrieron que la condensación era su método de elección. Sólo requiere la habilidad para calcular determinantes de 2×2.

Se requiere la siguiente terminología.

Definición Si A es una matriz de $n \times n$, con $n \geq 3$, el ***interior*** de A, denotado $\text{int}(A)$, es la matriz $(n-2) \times (n-2)$ que se obtiene al borrar el primer renglón, el último renglón, la primera columna y la última columna de A.

Se ilustrará el método de condensación para la matriz de 5×5

$$A = \begin{bmatrix} 2 & 3 & -1 & 2 & 0 \\ 1 & 2 & 3 & 1 & -4 \\ 2 & -1 & 2 & 1 & 1 \\ 3 & 1 & -1 & 2 & -2 \\ -4 & 1 & 0 & 1 & 2 \end{bmatrix}$$

Comience por hacer A_0 igual a la matriz de 6×6 cuyas entradas son todas 1. Luego, establezca $A_1 = A$. Es útil imaginar A_0 como la base de una pirámide con A_1 centrada encima de A_0. Agregue a la pirámide capas sucesivamente más pequeñas, hasta que llegar a una matriz de 1×1 en la parte superior, ésta contendrá det A. (Figura 4.9)

Esta viñeta se basa en el artículo "Lewis Carroll's Condensation Method for Evaluating Determinants" de Adrian Rice y Eve Torrence, en *Math Horizons*, noviembre de 2006, pp. 12-15. Para más detalles acerca del método de condensación, vea David M. Bressoud, *Proofs and Confirmations: The Story of the Alternating Sign Matrix Conjecture*, MAA Spectrum Series (Cambridge University Press, 1999).

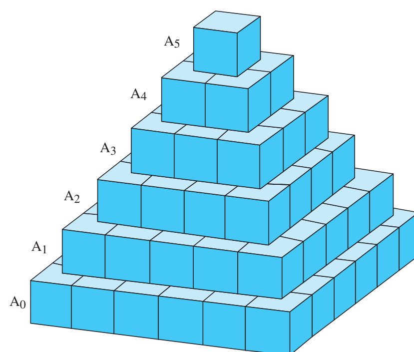


Figura 4.9

A continuación, "condense" A_1 en una matriz A'_2 de 4×4 cuyas entradas son los determinantes de todas las submatrices de 2×2 de A_1:

$$A'_2 = \begin{bmatrix} \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 3 & -1 \\ 2 & 3 \end{vmatrix} & \begin{vmatrix} -1 & 2 \\ 3 & 1 \end{vmatrix} & \begin{vmatrix} 2 & 0 \\ 1 & -4 \end{vmatrix} \\ \begin{vmatrix} 1 & 2 \\ 2 & -1 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ -1 & 2 \end{vmatrix} & \begin{vmatrix} 3 & 1 \\ 2 & 1 \end{vmatrix} & \begin{vmatrix} 1 & -4 \\ 1 & 1 \end{vmatrix} \\ \begin{vmatrix} 2 & -1 \\ 3 & 1 \end{vmatrix} & \begin{vmatrix} -1 & 2 \\ 1 & -1 \end{vmatrix} & \begin{vmatrix} 2 & 1 \\ -1 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 2 & -2 \end{vmatrix} \\ \begin{vmatrix} 3 & 1 \\ -4 & 1 \end{vmatrix} & \begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} & \begin{vmatrix} -1 & 2 \\ 0 & 1 \end{vmatrix} & \begin{vmatrix} 2 & -2 \\ 1 & 2 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 1 & 11 & -7 & -8 \\ -5 & 7 & 1 & 5 \\ 5 & -1 & 5 & -4 \\ 7 & 1 & -1 & 6 \end{bmatrix}$$

Ahora divida cada entrada de A'_2 por la correspondiente entrada de $\text{int}(A_0)$ para obtener la matriz A_2. Dado que A_0 es toda números 1, esto significa $A_2 = A'_2$.

Repita el procedimiento y construya A'_3 a partir de las submatrices de 2×2 de A_2 y luego divida cada entrada de A'_3 por la correspondiente entrada de $\text{int}(A_1)$ y continúe así. Se obtiene:

$$A'_3 = \begin{bmatrix} \begin{vmatrix} 1 & 11 \\ -5 & 7 \end{vmatrix} & \begin{vmatrix} 11 & -7 \\ 7 & 1 \end{vmatrix} & \begin{vmatrix} -7 & -8 \\ 1 & 5 \end{vmatrix} \\ \begin{vmatrix} -5 & 7 \\ 5 & -1 \end{vmatrix} & \begin{vmatrix} 7 & 1 \\ -1 & 5 \end{vmatrix} & \begin{vmatrix} 1 & 5 \\ 5 & -4 \end{vmatrix} \\ \begin{vmatrix} 5 & -1 \\ 7 & 1 \end{vmatrix} & \begin{vmatrix} -1 & 5 \\ 1 & -1 \end{vmatrix} & \begin{vmatrix} 5 & -4 \\ -1 & 6 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} 62 & 60 & -27 \\ -30 & 36 & -29 \\ 12 & -4 & 26 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} 62/2 & 60/3 & -27/1 \\ -30/-1 & 36/2 & -29/1 \\ 12/1 & -4/-1 & 26/2 \end{bmatrix} = \begin{bmatrix} 31 & 20 & -27 \\ 30 & 18 & -29 \\ 12 & 4 & 13 \end{bmatrix},$$

$$A'_4 = \begin{bmatrix} \begin{vmatrix} 31 & 20 \\ 30 & 18 \end{vmatrix} & \begin{vmatrix} 20 & -27 \\ 18 & -29 \end{vmatrix} \\ \begin{vmatrix} 30 & 18 \\ 12 & 4 \end{vmatrix} & \begin{vmatrix} 18 & -29 \\ 4 & 13 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -42 & -94 \\ -96 & 350 \end{bmatrix},$$

$$A_4 = \begin{bmatrix} -42/7 & -94/1 \\ -96/(-1) & 350/5 \end{bmatrix} = \begin{bmatrix} -6 & -94 \\ 96 & 70 \end{bmatrix},$$

$$A'_5 = \left[\begin{vmatrix} -6 & -94 \\ 96 & 70 \end{vmatrix} \right] = [8604],$$

$$A_5 = [8604/18] = [478]$$

Como puede comprobar por otros métodos, $A = 478$. En general, para una matriz A de $n \times n$, el método de condensación producirá una matriz A_n de 1×1 que contiene det A.

Claramente, el método se descompone si el interior de cualquiera de las A_i contiene un cero, pues entonces se trataría de dividir entre cero para construir A_{i+1}. Sin embargo, puede usar cuidadosamente las operaciones elementales con renglones y columnas para eliminar los ceros, de modo que pueda avanzar.

Exploración

Aplicaciones geométricas de los determinantes

Esta exploración revelará algunas de las sorprendentes aplicaciones de los determinantes a la geometría. En particular, se verá que los determinantes están estrechamente relacionados con fórmulas para área y volumen, y pueden usarse para producir las ecuaciones de líneas, planos y otros tipos de curvas. La mayoría de estas ideas surgieron cuando la teoría de los determinantes se desarrollaba como un tema por derecho propio.

El producto cruz

Recuerde de la Exploración: "El producto cruz" del capítulo 1, que el producto cruz de $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ es el vector $\mathbf{u} \times \mathbf{v}$ definido por

$$\mathbf{u} \times \mathbf{v} = \begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$$

Si este producto cruz se escribe como $(u_2v_3 - u_3v_2)\mathbf{e}_1 - (u_1v_3 - u_3v_1)\mathbf{e}_2 + (u_1v_2 - u_2v_1)\mathbf{e}_3$, donde $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ son los vectores base estándar, entonces se ve que la *forma* de esta fórmula es

$$\mathbf{u} \times \mathbf{v} = \det \begin{bmatrix} \mathbf{e}_1 & u_1 & v_1 \\ \mathbf{e}_2 & u_2 & v_2 \\ \mathbf{e}_3 & u_3 & v_3 \end{bmatrix}$$

si se expande a lo largo de la primera columna. (Desde luego, este no es un determinante propio, pues $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ son vectores, no escalares; sin embargo, ofrece una forma útil de recordar la fórmula un tanto complicada del producto cruz. También permite el uso de las propiedades de los determinantes para verificar algunas de las otras propiedades del producto cruz.)

Ahora se repasarán algunos de los ejercicios del capítulo 1.

1. Use la versión en determinantes del producto cruz para calcular $\mathbf{u} \times \mathbf{v}$.

(a) $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$ (b) $\mathbf{u} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

(c) $\mathbf{u} = \begin{bmatrix} -1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -4 \\ -6 \end{bmatrix}$ (d) $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

2. Si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$, y $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$, demuestre que

$$\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = \det \begin{bmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{bmatrix}$$

3. Use las propiedades de los determinantes (y el problema 2 anterior, si es necesario) para probar la propiedad dada del producto cruz.

(a) $\mathbf{v} \times \mathbf{u} = -(\mathbf{u} \times \mathbf{v})$ (b) $\mathbf{u} \times \mathbf{0} = \mathbf{0}$

(c) $\mathbf{u} \times \mathbf{u} = \mathbf{0}$ (d) $\mathbf{u} \times k\mathbf{v} = k(\mathbf{u} \times \mathbf{v})$

(e) $\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w}$ (f) $\mathbf{u} \cdot (\mathbf{u} \times \mathbf{v}) = 0$ y $\mathbf{v} \cdot (\mathbf{u} \times \mathbf{v}) = 0$

(g) $\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) = (\mathbf{u} \times \mathbf{v}) \cdot \mathbf{w}$ (la ***identidad del triple producto escalar***)

Área y volumen

Ahora se puede dar una interpretación geométrica de los determinantes de matrices de 2×2 y 3×3. Recuerde que si $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{R}^3$, entonces el área $\mathcal{A}$ del paralelogramo determinado por dichos vectores está dado por $\mathcal{A} = \|\mathbf{u} \times \mathbf{v}\|$. (Vea la Exploración: "El producto cruz" del capítulo 1.)

4. Sean $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. Demuestre que el área $\mathcal{A}$ del paralelogramo determinado por $\mathbf{u}$ y $\mathbf{v}$ está dada por

$$\mathcal{A} = \left| \det \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \end{bmatrix} \right|$$

[*Sugerencia:* escriba $\mathbf{u}$ y $\mathbf{v}$ como $\begin{bmatrix} u_1 \\ u_2 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} v_1 \\ v_2 \\ 0 \end{bmatrix}$.]

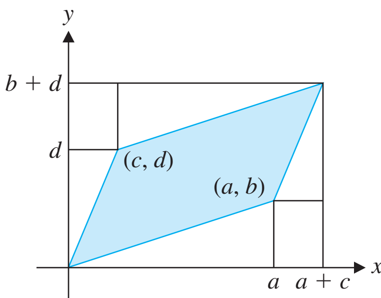


Figura 4.10

5. Obtenga geométricamente la fórmula de área del problema 4 y use la figura 4.10 como guía. [*Sugerencia:* reste áreas del rectángulo grande hasta que quede el paralelogramo.] En este caso, ¿de dónde proviene el signo del valor absoluto?

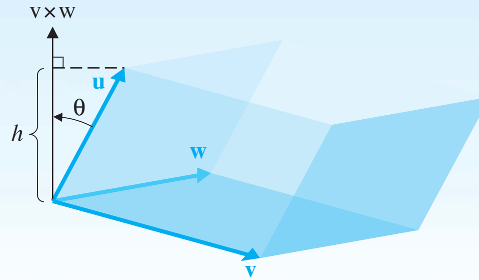


Figura 4.11

6. Encuentre el área del paralelogramo determinado por **u** y **v**.

$$\text{(a) } \mathbf{u} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -1 \\ 4 \end{bmatrix} \quad \text{(b) } \mathbf{u} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}$$

Al generalizar los problemas 4-6, considere un ***paralelepípedo***, un sólido tridimensional que recuerda un ladrillo "inclinado", cuyas seis caras son todas paralelogramos con caras opuestas paralelas y congruentes (figura 4.11). Su volumen está dado por el área de su base por su altura.

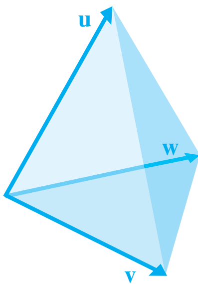


Figura 4.12

7. Demuestre que el volumen $\mathcal{V}$ del paralelepípedo determinado por **u**, **v** y **w** está dado por el valor absoluto del determinante de la matriz de 3×3 $[\mathbf{u} \quad \mathbf{v} \quad \mathbf{w}]$ con **u**, **v** y **w** como sus columnas. [*Sugerencia:* de la figura 4.11 puede ver que la altura h puede expresarse como $h = \|\mathbf{u}\| \cos\theta$, donde θ es el ángulo entre **u** y $\mathbf{v} \times \mathbf{w}$. Use este hecho para demostrar que $\mathcal{V} = |\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$ y aplique el resultado del problema 2.]

8. Demuestre que el volumen $\mathcal{V}$ del tetraedro determinado por **u**, **v** y **w** (figura 4.12) está dado por

$$\mathcal{V} = \tfrac{1}{6}|\mathbf{u} \cdot (\mathbf{v} \times \mathbf{w})|$$

[*Sugerencia:* de la geometría se sabe que el volumen de tal sólido es $\mathcal{V} = \frac{1}{3}$ (área de la base)(altura).]

Ahora vea estas interpretaciones geométricas desde un punto de vista transformacional. Sea A una matriz de 2×2 y sea P el paralelogramo determinado por los vectores **u** y **v**. Considere el efecto de la transformación matricial T_A sobre el área de P. Sea $T_A(P)$ el paralelogramo determinado por $T_A(\mathbf{u}) = A\mathbf{u}$ y $T_A(\mathbf{v}) = A\mathbf{v}$.

9. Demuestre que el área de $T_A(P)$ está dada por $|\det A|$(área de P).

10. Sea A una matriz de 3×3 y sea P el paralelepípedo determinado por los vectores **u**, **v** y **w**. Sea $T_A(P)$ que denota al paralelepípedo determinado por $T_A(\mathbf{u}) = A\mathbf{u}$, $T_A(\mathbf{v}) = A\mathbf{v}$ y $T_A(\mathbf{w}) = A\mathbf{w}$. Demuestre que el volumen de $T_A(P)$ está dado por $|\det A|$(volumen de P).

Los problemas anteriores ilustran que el determinante de una matriz captura lo que hace la correspondiente transformación matricial al área o volumen de las figuras sobre las que actúa. (Aunque sólo se consideraron ciertos tipos de figuras, el resultado es perfectamente general y puede hacerse riguroso. Eso no se hará aquí.)

Rectas y planos

Suponga que se dan dos puntos distintos (x_1, y_1) y (x_2, y_2) en el plano. Existe una recta única que pasa a través de dichos puntos y su ecuación es de la forma

$$ax + by + c = 0$$

Puesto que los dos puntos dados están sobre esta recta, sus coordenadas satisfacen esta ecuación. Por ende,

$$ax_1 + by_1 + c = 0$$
$$ax_2 + by_2 + c = 0$$

Las tres ecuaciones juntas pueden verse como un sistema de ecuaciones lineales con las variables a, b y c. Dado que *hay* una solución no trivial (es decir, la recta existe), la matriz de coeficientes

$$\begin{bmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix}$$

no puede ser invertible, por el teorema fundamental de las matrices invertibles. En consecuencia, su determinante debe ser cero, por el Teorema 4.6. Al expandir este determinante se obtiene la ecuación de la recta.

La ecuación de la recta que pasa por los puntos (x_1, y_1) y (x_2, y_2) está dada por

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0$$

11. Use el método descrito anteriormente para encontrar la ecuación de la recta que pasa por los puntos dados.

(a) $(2, 3)$ y $(-1, 0)$ (b) $(1, 2)$ y $(4, 3)$

12. Demuestre que los tres puntos (x_1, y_1), (x_2, y_2) y (x_3, y_3) son colineales (están en la misma recta) si y sólo si

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0$$

13. Demuestre que la ecuación del plano a través de los tres puntos no colineales (x_1, y_1, z_1), (x_2, y_2, z_2) y (x_3, y_3, z_3) está dada por

$$\begin{vmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{vmatrix} = 0$$

¿Qué sucede si los tres puntos *son* colineales? [*Sugerencia:* explique qué sucede cuando se usa reducción por renglón para evaluar el determinante.]

14. Demuestre que los cuatro puntos (x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3) y (x_4, y_4, z_4) son coplanares (están en el mismo plano) si y sólo si

$$\begin{vmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \\ x_4 & y_4 & z_4 & 1 \end{vmatrix} = 0$$

Ajuste de curvas

Cuando los datos que surgen de la experimentación toman la forma de puntos (x, y) que pueden graficarse en el plano, con frecuencia es de interés encontrar una relación entre las variables x y y. Idealmente se querría encontrar una función cuya gráfica pase a través de todos los puntos. En ocasiones todo lo que se quiere es una aproximación (vea la sección 7.3), pero en ciertas situaciones también son posibles resultados exactos.

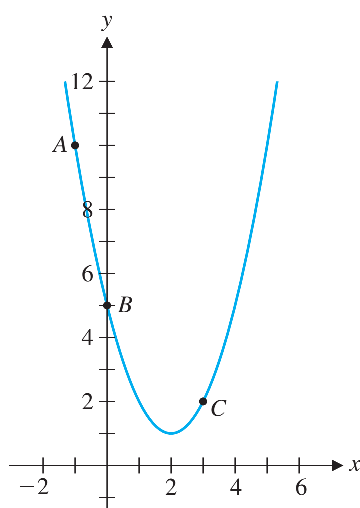


Figura 4.13

15. De la figura 4.13 parece como si pudiera encontrarse una parábola que pasa a través de los puntos $A(-1, 10)$, $B(0, 5)$ y $C(3, 2)$. La ecuación de tal parábola es de la forma $y = a + bx + cx^2$. Al sustituir los puntos dados en esta ecuación, establezca un sistema de tres ecuaciones lineales con las variables a, b y c. *Sin resolver el sistema*, use el Teorema 4.6 para argumentar que debe tener una solución única. Luego resuelva el sistema para encontrar la ecuación de la parábola en la figura 4.13.

16. Use el método del problema 15 para encontrar los polinomios de grado cuando mucho 2 que pasen a través de los siguientes conjuntos de puntos.

(a) $A(1, -1), B(2, 4), C(3, 3)$ (b) $A(-1, -3), B(1, -1), C(3, 1)$

17. Al generalizar los problemas 15 y 16, suponga que a_1, a_2 y a_3 son distintos números reales. Para cualquier número real b_1, b_2 y b_3, se quiere demostrar que existe una cuadrática única con ecuación de la forma $y = a + bx + cx^2$ que pasa a través de los puntos (a_1, b_1), (a_2, b_2) y (a_3, b_3). Haga esto para demostrar que la matriz de coeficientes del sistema lineal asociado tiene el determinante

$$\begin{vmatrix} 1 & a_1 & a_1^2 \\ 1 & a_2 & a_2^2 \\ 1 & a_3 & a_3^2 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1)(a_3 - a_2)$$

que necesariamente es distinto de cero. (¿Por qué?)

18. Sean a_1, a_2, a_3, y a_4 distintos números reales. Demuestre que

$$\begin{vmatrix} 1 & a_1 & a_1^2 & a_1^3 \\ 1 & a_2 & a_2^2 & a_2^3 \\ 1 & a_3 & a_3^2 & a_3^3 \\ 1 & a_4 & a_4^2 & a_4^3 \end{vmatrix} = (a_2 - a_1)(a_3 - a_1)(a_4 - a_1)(a_3 - a_2)(a_4 - a_2)(a_4 - a_3) \neq 0$$

Para cualesquiera números reales b_1, b_2, b_3 y b_4, use ese resultado para probar que existe una cúbica única con ecuación $y = a + bx + cx^2 + dx^3$ que pasa a través de los cuatro puntos (a_1, b_1), (a_2, b_2), (a_3, b_3) y (a_4, b_4). (*No* resuelva realmente para a, b, c y d.)

19. Sean $a_1, a_2, \ldots, a_n$ n números reales. Demuestre que

$$\begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{n-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{n-1} \\ 1 & a_3 & a_3^2 & \cdots & a_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i)$$

donde $\Pi_{1 \leq i < j \leq n} (a_j - a_i)$ significa el producto de todos los términos de la forma $(a_j - a_i)$, donde $i < j$ y tanto i como j están entre 1 y n. [El determinante de una matriz de esta forma (o su transpuesta) se llama ***determinante de Vandermonde***, en honor del matemático francés A. T. Vandermonde (1735-1796).]

Deduzca que, para cualesquiera n puntos en el plano cuyas coordenadas x son todas distintas, existe un polinomio único de grado $n - 1$ cuyo gráfico pasa a través de los puntos dados.

4.3 Eigenvalores y eigenvectores de matrices de $n \times n$

Ahora que se definió el determinante de una matriz de $n \times n$, puede continuar la discusión de los eigenvalores y los eigenvectores en un contexto general. Recuerde de la sección 4.1 que λ es un eigenvalor de A si y sólo si $A - \lambda I$ es no invertible. Por el Teorema 4.6, esto es verdadero si y sólo si $\det(A - \lambda I) = 0$. Para resumir:

Los eigenvalores de una matriz cuadrada A son precisamente las soluciones λ de la ecuación

$$\det(A - \lambda I) = 0$$

Cuando se expande $\det(A - \lambda I)$, se obtiene un polinomio en λ, llamado ***polinomio característico*** de A. La ecuación $\det(A - \lambda I) = 0$ se llama ***ecuación característica*** de A. Por ejemplo, si $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, su polinomio característico es

$$\det(A - \lambda I) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc)$$

Si A es de $n \times n$, su polinomio característico será de grado n. De acuerdo con el teorema fundamental del álgebra (vea el Apéndice D), un polinomio de grado n con coeficientes reales o complejos tiene cuando mucho n raíces distintas. Al aplicar este hecho al polinomio característico, se ve que una matriz de $n \times n$ con entradas reales o complejas tiene cuando mucho n eigenvalores distintos.

A continuación se resumirá el procedimiento que se seguirá (por ahora) para encontrar los eigenvalores y eigenvectores (eigenespacios) de una matriz.

Sea A una matriz de $n \times n$.

1. Calcule el polinomio característico $\det(A - \lambda I)$ de A.
2. Encuentre los eigenvalores de A al resolver la ecuación característica $\det(A - \lambda I) = 0$ para λ.
3. Para cada eigenvalor λ, encuentre el espacio nulo de la matriz $A - \lambda I$. Este es el eigenespacio E_λ, los vectores distintos de cero de los cuales surgen los eigenvectores de A correspondientes a λ.
4. Encuentre una base para cada eigenespacio.

Ejemplo 4.18

Encuentre los eigenvalores y los correspondientes eigenespacios de

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$$

Solución Se tiene el procedimiento resaltado anteriormente. El polinomio característico es

$$\begin{aligned}\det(A - \lambda I) &= \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 2 & -5 & 4-\lambda \end{vmatrix} \\ &= -\lambda\begin{vmatrix} -\lambda & 1 \\ -5 & 4-\lambda \end{vmatrix} - \begin{vmatrix} 0 & 1 \\ 2 & 4-\lambda \end{vmatrix} \\ &= -\lambda(\lambda^2 - 4\lambda + 5) - (-2) \\ &= -\lambda^3 + 4\lambda^2 - 5\lambda + 2\end{aligned}$$

Para encontrar los eigenvalores es necesario resolver la ecuación característica $\det(A - \lambda I) = 0$ para λ. El polinomio característico se factoriza como $-(\lambda - 1)^2(\lambda - 2)$. (El teorema del factor es útil aquí; vea el Apéndice D.) Por ende, la ecuación característica es $-(\lambda - 1)^2(\lambda - 2) = 0$, que claramente tiene soluciones $\lambda = 1$ y $\lambda = 2$. Dado que $\lambda = 1$ es una raíz múltiplo y $\lambda = 2$ es una raíz simple, etiquételas $\lambda_1 = \lambda_2 = 1$ y $\lambda_3 = 2$.

Para encontrar los eigenvectores correspondientes a $\lambda_1 = \lambda_2 = 1$, se encuentra el espacio nulo de

$$A - 1I = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & -5 & 4-1 \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 2 & -5 & 3 \end{bmatrix}$$

La reducción por renglones produce

$$[A - I \mid 0] = \left[\begin{array}{ccc|c} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 2 & -5 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$



(Se sabía por anticipado que *debía* tener al menos un renglón cero. ¿Por qué?) En consecuencia, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ está en el eigenespacio E_1 si y sólo si $x_1 - x_3 = 0$ y $x_2 - x_3 = 0$. Al establecer la variable libre $x_3 = t$, se ve que $x_1 = t$ y $x_2 = t$, de donde se tiene que

$$E_1 = \left\{\begin{bmatrix} t \\ t \\ t \end{bmatrix}\right\} = \left\{t\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\} = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right)$$

Para encontrar los eigenvectores correspondientes a $\lambda_3 = 2$, se encuentra el espacio nulo de $A - 2I$ mediante reducción por renglones:

$$[A - 2I \mid 0] = \left[\begin{array}{ccc|c} -2 & 1 & 0 & 0 \\ 0 & -2 & 1 & 0 \\ 2 & -5 & 2 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -\frac{1}{4} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

De modo que $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ está en el eigenespacio E_2 si y sólo si $x_1 = \frac{1}{4}x_3$ y $x_2 = \frac{1}{2}x_3$. Al establecer la variable libre $x_3 = t$, se tiene

$$E_2 = \left\{\begin{bmatrix} \frac{1}{4}t \\ \frac{1}{2}t \\ t \end{bmatrix}\right\} = \left\{t\begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ 1 \end{bmatrix}\right\} = \text{gen}\left(\begin{bmatrix} \frac{1}{4} \\ \frac{1}{2} \\ 1 \end{bmatrix}\right) = \text{gen}\left(\begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}\right)$$

donde se eliminan los denominadores en la base al multiplicar por el mínimo común denominador 4. (¿Por qué esto es permisible?)

Comentario Note que en el ejemplo 4.18 A es una matriz de 3×3 pero sólo tiene dos eigenvalores distintivos. Sin embargo, si se cuentan multiplicidades, A tiene exactamente tres eigenvalores ($\lambda = 1$ dos veces y $\lambda = 2$ una vez). Esto es lo que garantiza el teorema fundamental del álgebra. La ***multiplicidad algebraica*** de un eigenvalor es como su multiplicidad como una raíz de la ecuación característica. Por tanto, $\lambda = 1$ tiene multiplicidad algebraica 2 y $\lambda = 2$ tiene multiplicidad algebraica 1.

A continuación, note que cada eigenespacio tiene una base que consiste de justo un vector. En otras palabras, $\dim E_1 = \dim E_2 = 1$. La ***multiplicidad geométrica*** de un eigenvalor λ se define como $\dim E_\lambda$, la dimensión de su correspondiente eigenespacio. Como verá en la sección 4.4, es importante una comparación de estas dos nociones de multiplicidad.

Ejemplo 4.19 Encuentre los eigenvalores y los correspondientes eigenespacios de

$$A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$$

Solución La ecuación característica es

$$0 = \det(A - \lambda I) = \begin{vmatrix} -1-\lambda & 0 & 1 \\ 3 & -\lambda & -3 \\ 1 & 0 & -1-\lambda \end{vmatrix} = -\lambda \begin{vmatrix} -1-\lambda & 1 \\ 1 & -1-\lambda \end{vmatrix}$$
$$= -\lambda(\lambda^2 + 2\lambda) = -\lambda^2(\lambda + 2)$$

Por tanto, los eigenvalores son $\lambda_1 = \lambda_2 = 0$ y $\lambda_3 = -2$. En consecuencia, el eigenvalor 0 tiene multiplicidad algebraica 2 y el eigenvalor -2 tiene multiplicidad algebraica 1.

Para $\lambda_1 = \lambda_2 = 0$, se calcula

$$[A - 0I\,|\,\mathbf{0}] = [A\,|\,\mathbf{0}] = \left[\begin{array}{ccc|c} -1 & 0 & 1 & 0 \\ 3 & 0 & -3 & 0 \\ 1 & 0 & -1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

de donde se tiene que un eigenvector $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ en E_0 satisface $x_1 = x_3$. Por ende, tanto x_2 como x_3 son libres. Al hacer $x_2 = s$ y $x_3 = t$, se tiene

$$E_0 = \left\{ \begin{bmatrix} t \\ s \\ t \end{bmatrix} \right\} = \left\{ s\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\} = \text{gen}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right)$$

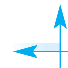
Para $\lambda_3 = -2$,

$$[A - (-2)I\,|\,\mathbf{0}] = [A + 2I\,|\,\mathbf{0}] = \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 3 & 2 & -3 & 0 \\ 1 & 0 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

de modo que $x_3 = t$ es libre y $x_1 = -x_3 = -t$ y $x_2 = 3x_3 = 3t$. En consecuencia,

$$E_{-2} = \left\{ \begin{bmatrix} -t \\ 3t \\ t \end{bmatrix} \right\} = \left\{ t \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right\} = \text{gen} \left(\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix} \right)$$

Se tiene que $\lambda_1 = \lambda_2 = 0$ posee multiplicidad geométrica 2 y $\lambda_3 = -2$ tiene multiplicidad geométrica 1. (Note que la multiplicidad algebraica es igual a la multiplicidad geométrica para cada eigenvalor.)



En algunas situaciones, los eigenvalores de una matriz son muy fáciles de encontrar. Si A es una matriz triangular, entonces también lo es $A - \lambda I$, y el Teorema 4.2 dice que $\det(A - \lambda I)$ es justo el producto de las entradas diagonales. Esto implica que la ecuación característica de una matriz triangular es

$$(a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = 0$$

de donde se tiene inmediatamente que los eigenvalores son $\lambda_1 = a_{11}$, $\lambda_2 = a_{22}$, . . . , $\lambda_n = a_{nn}$. Este resultado se resume como un teorema y se ilustra con un ejemplo.

Teorema 4.15 Los eigenvalores de una matriz triangular son las entradas en su diagonal principal.

Ejemplo 4.20 Los eigenvalores de

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 3 & 0 & 3 & 0 \\ 5 & 7 & 4 & -2 \end{bmatrix}$$

son $\lambda_1 = 2$, $\lambda_2 = 1$, $\lambda_3 = 3$ y $\lambda_4 = -2$, por el Teorema 4.15. De hecho, el polinomio característico es justo $(2 - \lambda)(1 - \lambda)(3 - \lambda)(-2 - \lambda)$.



Note que las matrices diagonales son un caso especial del Teorema 4.15. De hecho, una matriz diagonal es tanto triangular superior como inferior.

Los eigenvalores capturan información mucho más importante acerca del comportamiento de una matriz. Una vez se conocen los eigenvalores de una matriz, puede deducir una gran cantidad de cosas sin hacer más trabajo. El siguiente teorema es uno de los más importantes en este aspecto.

Teorema 4.16 Una matriz cuadrada A es invertible si y sólo si 0 *no* es un eigenvalor de A.

Demostración Sea A una matriz cuadrada. Por el Teorema 4.6, A es invertible si y sólo si $\det A \neq 0$. Pero $\det A \neq 0$ es equivalente a $\det(A - 0I) \neq 0$, lo cual dice que 0 no es una raíz de la ecuación característica de A (esto es, 0 no es un eigenvalor de A).

Ahora puede extender el teorema fundamental de las matrices invertibles para incluir los resultados que se han probado en este capítulo.

Teorema 4.17 **El teorema fundamental de las matrices invertibles: versión 3**

Sea A una matriz de $n \times n$. Los siguientes enunciados son equivalentes:

a. A es invertible.
b. $A\mathbf{x} = \mathbf{b}$ tiene una solución única para todo $\mathbf{b}$ en $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial.
d. La forma escalonada reducida por renglón de A es I_n.
e. A es un producto de matrices elementales.
f. $\text{rank}(A) = n$
g. $\text{nulidad}(A) = 0$
h. Los vectores columna de A son linealmente independientes.
i. Los vectores columna de A generan $\mathbb{R}^n$.
j. Los vectores columna de A forman una base para $\mathbb{R}^n$.
k. Los vectores renglón de A son linealmente independientes.
l. Los vectores renglón de A generan $\mathbb{R}^n$.
m. Los vectores renglón de A forman una base para $\mathbb{R}^n$.
n. $\det A \neq 0$
o. 0 no es un eigenvalor de A.

Demostración La equivalencia (a) $\Leftrightarrow$ (n) es el Teorema 4.6, y se demostró (a) $\Leftrightarrow$ (o) en el Teorema 4.16.

Existen fórmulas amigables para los eigenvalores de las potencias y la inversa de una matriz.

Teorema 4.18 Sea A una matriz cuadrada con eigenvalor λ y eigenvector correspondiente $\mathbf{x}$.

a. Para cualquier entero positivo n, λ^n es un eigenvalor de A^n con eigenvector correspondiente $\mathbf{x}$.
b. Si A es invertible, entonces $1/\lambda$ es un eigenvalor de A^{-1} con eigenvector correspondiente $\mathbf{x}$.
c. Si A es invertible, entonces, para cualquier entero n, λ^n es un eigenvalor de A^n con eigenvector correspondiente $\mathbf{x}$.

Demostración Se tiene $A\mathbf{x} = \lambda\mathbf{x}$.

(a) Se procede por inducción sobre n. Para $n = 1$, el resultado es el que se acaba de proporcionar. Suponga que el resultado es verdadero para $n = k$. Esto es, suponga que para algún entero positivo k, $A^k\mathbf{x} = \lambda^k\mathbf{x}$. Se debe demostrar el resultado para $n = k + 1$. Pero

$$A^{k+1}\mathbf{x} = A(A^k\mathbf{x}) = A(\lambda^k\mathbf{x})$$

por la hipótesis de inducción. Al usar la propiedad (d) del Teorema 3.3, se tiene

$$A(\lambda^k\mathbf{x}) = \lambda^k(A\mathbf{x}) = \lambda^k(\lambda\mathbf{x}) = \lambda^{k+1}\mathbf{x}$$

Por tanto, $A^{k+1}\mathbf{x} = \lambda^{k+1}\mathbf{x}$, como se requiere. Por inducción, el resultado es verdadero para todos los enteros $n \geq 1$.

(b) En el ejercicio 13 se le pide demostrar esta propiedad.

(c) En el ejercicio 14 se le pide demostrar esta propiedad.

El siguiente ejemplo muestra una aplicación de este teorema.

Ejemplo 4.21 Calcule $\begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}^{10} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

Solución Sean $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$ y $\mathbf{x} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$; entonces lo que se quiere encontrar es $A^{10}\mathbf{x}$.

Los eigenvalores de A son $\lambda_1 = -1$ y $\lambda_2 = 2$, con eigenvectores correspondientes $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Esto es,

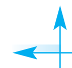
$$A\mathbf{v}_1 = -\mathbf{v}_1 \quad \text{y} \quad A\mathbf{v}_2 = 2\mathbf{v}_2$$


(Compruebe esto.) Dado que $\{\mathbf{v}_1, \mathbf{v}_2\}$ forma una base para $\mathbb{R}^2$ (¿por qué?), puede escribir $\mathbf{x}$ como una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$. De hecho, como se comprueba fácilmente, $\mathbf{x} = 3\mathbf{v}_1 + 2\mathbf{v}_2$.

En consecuencia, al usar el Teorema 4.18(a), se tiene

$$\begin{aligned} A^{10}\mathbf{x} &= A^{10}(3\mathbf{v}_1 + 2\mathbf{v}_2) = 3(A^{10}\mathbf{v}_1) + 2(A^{10}\mathbf{v}_2) \\ &= 3(\lambda_1^{10})\mathbf{v}_1 + 2(\lambda_2^{10})\mathbf{v}_2 \\ &= 3(-1)^{10}\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 2(2^{10})\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 + 2^{11} \\ -3 + 2^{12} \end{bmatrix} = \begin{bmatrix} 2051 \\ 4093 \end{bmatrix} \end{aligned}$$

Ciertamente, esto es mucho más sencillo que primero calcular A^{10}; de hecho, ¡no existen multiplicaciones matriciales!



Cuando se puede usar, el método del ejemplo 4.21 es bastante general. Se le resume como el siguiente teorema, que se le pedirá demostrar en el ejercicio 42.

Teorema 4.19 Suponga que la matriz A de $n \times n$ tiene eigenvectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ con correspondientes eigenvalores $\lambda_1, \lambda_2, \ldots, \lambda_m$. Si $\mathbf{x}$ es un vector en $\mathbb{R}^n$ que puede expresarse como una combinación lineal de estos eigenvectores, por decir,

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_m\mathbf{v}_m$$

entonces, para cualquier entero k,

$$A^k\mathbf{x} = c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \cdots + c_m\lambda_m^k\mathbf{v}_m$$

Advertencia El truco aquí es el "si" en la segunda oración. Absolutamente no hay garantía de que tal combinación lineal sea posible. La mejor situación posible sería si hubiera una base de $\mathbb{R}^n$ que consistiera de eigenvectores de A; en la siguiente sección se explorará esta posibilidad. Sin embargo, como un paso en dicha dirección, se tiene el siguiente teorema, que afirma que los eigenvectores correspondientes a *distintos* eigenvalores son linealmente independientes.

Teorema 4.20 Sea A una matriz de $n \times n$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_m$ distintos eigenvalores de A con sus correspondientes eigenvectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$. Entonces $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente independientes.

Demostración La demostración es indirecta. Se supondrá que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente *dependientes* y se demostrará que esta suposición conduce a una contradicción.

Si $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes, entonces uno de dichos vectores debe poder expresarse como una combinación lineal de los anteriores. Sea $\mathbf{v}_{k+1}$ el primero de los vectores $\mathbf{v}_i$ que pueden expresarse de esa forma. En otras palabras, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ son linealmente independientes, pero existen escalares $c_1, c_2, \ldots, c_k$ tales que

$$\mathbf{v}_{k+1} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k \tag{1}$$

Al multiplicar por la izquierda ambos lados de la ecuación (1) por A y con el hecho de que $A\mathbf{v}_i\ \lambda_i\mathbf{v}_i$ para cada i, se tiene

$$\begin{aligned} \lambda_{k+1}\mathbf{v}_{k+1} = A\mathbf{v}_{k+1} &= A(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) \\ &= c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 + \cdots + c_kA\mathbf{v}_k \\ &= c_1\lambda_1\mathbf{v}_1 + c_2\lambda_2\mathbf{v}_2 + \cdots + c_k\lambda_k\mathbf{v}_k \end{aligned} \tag{2}$$

Ahora multiplique ambos lados de la ecuación (1) por λ_{k+1} para obtener

$$\lambda_{k+1}\mathbf{v}_{k+1} = c_1\lambda_{k+1}\mathbf{v}_1 + c_2\lambda_{k+1}\mathbf{v}_2 + \cdots + c_k\lambda_{k+1}\mathbf{v}_k \tag{3}$$

Cuando se resta la ecuación (3) de la ecuación (2), se obtiene

$$\mathbf{0} = c_1(\lambda_1 - \lambda_{k+1})\mathbf{v}_1 + c_2(\lambda_2 - \lambda_{k+1})\mathbf{v}_2 + \cdots + c_k(\lambda_k - \lambda_{k+1})\mathbf{v}_k$$

La independencia lineal de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ implica que

$$c_1(\lambda_1 - \lambda_{k+1}) = c_2(\lambda_2 - \lambda_{k+1}) = \cdots = c_k(\lambda_k - \lambda_{k+1}) = 0$$

Dado que los eigenvalores λ_i son todos diferentes, los términos entre paréntesis $(\lambda_i - \lambda_{k+1})$, $i = 1, \ldots, k$, son todos distintos de cero. Por tanto, $c_1 = c_2 = \cdots = c_k = 0$. Esto implica que

$$\mathbf{v}_{k+1} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_k = \mathbf{0}$$

lo que es imposible, pues el eigenvector $\mathbf{v}_{k+1}$ no puede ser cero. En consecuencia, se tiene una contradicción, lo cual significa que la suposición de que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ son linealmente dependientes es falsa. Se tiene que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ deben ser linealmente independientes.

Ejercicios 4.3

En los ejercicios 1-12, calcule (a) el polinomio característico de A, (b) los eigenvalores de A, (c) una base para cada eigenespacio de A y (d) la multiplicidad algebraica y geométrica de cada eigenvalor.

1. $A = \begin{bmatrix} 1 & 3 \\ -2 & 6 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & -9 \\ 1 & -5 \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

5. $A = \begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ -1 & 1 & 1 \end{bmatrix}$

7. $A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ -1 & 0 & 2 \end{bmatrix}$

8. $A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & -2 & 2 \\ 3 & 0 & 1 \end{bmatrix}$

9. $A = \begin{bmatrix} 3 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

10. $A = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 2 \end{bmatrix}$

11. $A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 3 & 0 \\ -2 & 1 & 2 & -1 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 4 & 1 & 0 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 2 & 0 \end{bmatrix}$

13. Demuestre el Teorema 4.18(b).

14. Demuestre el Teorema 4.18(c). [*Sugerencia:* combine las demostraciones de los incisos (a) y (b) y vea el cuarto Comentario que sigue al Teorema 3.9 (página 175).]

En los ejercicios 15 y 16, A es una matriz de 2×2 con eigenvectores $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ *y* $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *que corresponden a eigenvalores* $\lambda_1 = \frac{1}{2}$ *y* $\lambda_2 = 2$, *respectivamente, y* $\mathbf{x} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$.

15. Encuentre $A^{10}\mathbf{x}$.

16. Encuentre $A^k\mathbf{x}$. ¿Qué sucede conforme k se vuelve grande (esto es, $k \to \infty$)?

En los ejercicios 17 y 18, A es una matriz de 3×3 con eigenvectores $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ *y* $\mathbf{v}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ *correspondientes a eigenvalores* $\lambda_1 = -\frac{1}{3}$, $\lambda_2 = \frac{1}{3}$ *y* $\lambda_3 = 1$, *respectivamente, y* $\mathbf{x} = \begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix}$.

17. Encuentre $A^{20}\mathbf{x}$.

18. Encuentre $A^k\mathbf{x}$. ¿Qué sucede conforme k se vuelve grande (esto es, $k \to \infty$)?

19. **(a)** Demuestre que para cualquier matriz cuadrada A, A^T y A tienen el mismo polinomio característico y en consecuencia los mismos eigenvalores.
(b) Dé un ejemplo de una matriz A de 2×2 para la cual A^T y A tengan diferentes eigenespacios.

20. Sea A una matriz nilpotente (esto es: $A^m = O$ para algún $m > 1$). Demuestre que $\lambda = 0$ es el único eigenvalor de A.

21. Sea A una matriz idempotente (esto es, $A^2 = A$). Demuestre que $\lambda = 0$ y $\lambda = 1$ son los únicos eigenvalores posibles de A.

22. Si $\mathbf{v}$ es un eigenvector de A con su correspondiente eigenvalor λ y c es un escalar, demuestre que $\mathbf{v}$ es un eigenvector de $A - cI$ con su eigenvalor correspondiente $\lambda - c$.

23. **(a)** Encuentre los eigenvalores y eigenespacios de

$$A = \begin{bmatrix} 3 & 2 \\ 5 & 0 \end{bmatrix}$$

(b) Con el Teorema 4.18 y el ejercicio 22, encuentre los eigenvalores y eigenespacios de A^{-1}, $A - 2I$ y $A + 2I$.

24. Sean A y B matrices de $n \times n$ con eigenvalores λ y μ, respectivamente.
(a) Ofrezca un ejemplo para demostrar que $\lambda + \mu$ no necesita ser un eigenvalor de $A + B$.
(b) Ofrezca un ejemplo para demostrar que $\lambda\mu$ no necesita ser un eigenvalor de AB.
(c) Suponga que λ y μ corresponden al *mismo* eigenvector $\mathbf{x}$. Demuestre que, en este caso, $\lambda + \mu$ es un eigenvalor de $A + B$ y $\lambda\mu$ es un eigenvalor de AB.

25. Si A y B son dos matrices equivalentes por renglon, ¿necesariamente tienen los mismos eigenvalores? Demuestre que lo son u ofrezca un contraejemplo.

Sea $p(x)$ el polinomio

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$$

La ***matriz acompañante*** *de* $p(x)$ es la matriz de $n \times n$

$$C(p) = \begin{bmatrix} -a_{n-1} & -a_{n-2} & \cdots & -a_1 & -a_0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad (4)$$

26. Encuentre la matriz acompañante de $p(x) = x^2 - 7x + 12$ y luego encuentre el polinomio característico de $C(p)$.

27. Encuentre la matriz acompañante de $p(x) = x^3 + 3x^2 - 4x + 12$ y luego encuentre el polinomio característico de $C(p)$.

28. **(a)** Demuestre que la matriz acompañante $C(p)$ de $p(x) = x^2 + ax + b$ tiene polinomio característico $\lambda^2 + a\lambda + b$.
(b) Demuestre que si λ es un eigenvalor de la matriz acompañante $C(p)$ en el inciso (a), entonces $\begin{bmatrix} \lambda \\ 1 \end{bmatrix}$ es un eigenvector de $C(p)$ correspondiente a λ.

29. **(a)** Demuestre que la matriz acompañante $C(p)$ de $p(x) = x^3 + ax^2 + bx + c$ tiene polinomio característico $-(\lambda^3 + a\lambda^2 + b\lambda + c)$.
(b) Demuestre que, si λ es un eigenvalor de la matriz acompañante $C(p)$ en el inciso (a), entonces $\begin{bmatrix} \lambda^2 \\ \lambda \\ 1 \end{bmatrix}$ es un eigenvector de $C(p)$ correspondiente a λ.

30. Construya una matriz no triangular de 2×2 con eigenvalores 2 y 5. [*Sugerencia:* use el ejercicio 28.]

31. Construya una matriz no triangular de 3×3 con eigenvalores -2, 1 y 3. [*Sugerencia:* use el ejercicio 29.]

32. **(a)** Use inducción matemática para demostrar que, para $n \geq 2$, la matriz acompañante $C(p)$ de $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ tiene polinomio característico $(-1)^n p(\lambda)$. [*Sugerencia:* desarrolle por cofactores a lo largo de la última columna. Puede descubrir que es útil introducir el polinomio $q(x) = (p(x) - a_0)/x$.]

(b) Demuestre que si λ es un eigenvalor de la matriz acompañante $C(p)$ en la ecuación (4), entonces un eigenvector correspondiente a λ está dado por

$$\begin{bmatrix} \lambda^{n-1} \\ \lambda^{n-2} \\ \vdots \\ \lambda \\ 1 \end{bmatrix}$$

Si $p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ y A es una matriz cuadrada, puede definir una matriz cuadrada p(A) mediante

$$p(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1A + a_0I$$

*Un importante teorema del álgebra lineal avanzada dice que si $c_A(\lambda)$ es el polinomio característico de la matriz A, entonces $c_A(A) = O$ (en palabras, toda matriz satisface su ecuación característica). Este es el famoso **teorema de Cayley-Hamilton**, llamado así en honor de Arthur Cayley (1821-1895), y de sir William Rowan Hamilton. Cayley demostró este teorema en 1858. Hamilton lo descubrió, independientemente, en su trabajo acerca de los cuaterniones, una generalización de los números complejos.*

33. Verifique el teorema de Cayley-Hamilton para

$A = \begin{bmatrix} 1 & -1 \\ 2 & 3 \end{bmatrix}$. Esto es, encuentre el polinomio característico $c_A(\lambda)$ de A y demuestre que $c_A(A) = O$.

34. Verifique el teorema de Cayley-Hamilton para

$$A = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}.$$

El teorema de Cayley-Hamilton puede usarse para calcular potencias e inversos de matrices. Por ejemplo, si A es una matriz de 2×2 con polinomio característico $c_A(\lambda) = \lambda^2 + a\lambda + b$, entonces $A^2 + aA + bI = O$, de modo que

$$A^2 = -aA - bI$$

y

$$\begin{aligned} A^3 = AA^2 &= A(-aA - bI) \\ &= -aA^2 - bA \\ &= -a(-aA - bI) - bA \\ &= (a^2 - b)A + abI \end{aligned}$$

Es fácil ver que, al continuar de esta forma, puede expresar cualquier potencia positiva de A como una combinación lineal de I y A. A partir de $A^2 + aA + bI = O$, también se obtiene $A(A + aI) = -bI$, de modo que

$$A^{-1} = -\frac{1}{b}A - \frac{a}{b}I$$

siempre que $b \neq 0$.

35. Para la matriz A del ejercicio 33, use el teorema de Cayley-Hamilton para calcular, A^2, A^3 y A^4 al expresar cada una como una combinación lineal de I y A.

36. Para la matriz A del ejercicio 34, use el teorema de Cayley-Hamilton para calcular A^3 y A^4 al expresar cada una como una combinación lineal de I, A y A^2.

37. Para la matriz A del ejercicio 33, use el teorema de Cayley-Hamilton para calcular A^{-1} y A^{-2} al expresar cada una como una combinación lineal de I y A.

38. Para la matriz A del ejercicio 34, use el teorema de Cayley-Hamilton para calcular A^{-1} y A^{-2} al expresar cada una como una combinación lineal de I, A y A^2.

39. Demuestre que si la matriz cuadrada A puede partirse como

$$A = \left[\begin{array}{c|c} P & Q \\ \hline O & S \end{array}\right]$$

donde P y S son matrices cuadradas, entonces el polinomio característico de A es $c_A(\lambda) = c_P(\lambda)c_S(\lambda)$. [*Sugerencia:* use el ejercicio 69 de la sección 4.2.]

40. Sea $\lambda_1, \lambda_2, \ldots, \lambda_n$ un conjunto completo de eigenvalores (incluidas repeticiones) de la matriz A de $n \times n$. Demuestre que

$$\det(A) = \lambda_1\lambda_2\cdots\lambda_n \quad \text{y}$$
$$\operatorname{tr}(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$$

[*Sugerencia:* el polinomio característico de A se factoriza como

$$\det(A - \lambda I) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n)$$

Encuentre el término constante y el coeficiente de λ^{n-1} en los lados izquierdo y derecho de esta ecuación.]

41. Sean A y B matrices de $n \times n$. Demuestre que la suma de *todos* los eigenvalores de $A + B$ es la suma de todos los eigenvalores de A y B individualmente. Pruebe que el producto de *todos* los eigenvalores de AB es el producto de todos los eigenvalores de A y B individualmente. (Compare este ejercicio con el ejercicio 24.)

42. Pruebe el Teorema 4.19.

4.4 Semejanza y diagonalización

Como vio en la última sección, las matrices triangulares y diagonales son agradables en el sentido de que sus eigenvalores se muestran con transparencia. Sería agradable si pudiera relacionarse una matriz cuadrada dada con una triangular o diagonal en tal forma que tuvieran exactamente los mismos eigenvalores. Desde luego, ya conoce un procedimiento para convertir una matriz cuadrada en forma triangular, a saber, la eliminación gaussiana. Por desgracia, este proceso no conserva los eigenvalores de la matriz. En esta sección se considera un tipo diferente de transformación de una matriz que se comporta bien con respecto a los eigenvalores.

Matrices semejantes

Definición Sean A y B matrices de $n \times n$. Se dice que A ***es semejante a*** B si existe una matriz invertible P de $n \times n$ tal que $P^{-1}AP = B$. Si A es semejante a B, se escribe $A \sim B$.

Comentarios

- Si $A \sim B$, puede escribir, de modo equivalente, que $A = PBP^{-1}$ o $AP = PB$.
- La semejanza es una *relación* entre matrices en el mismo sentido que "menor que o igual a" es una relación entre enteros. Note que en la definición hay una *dirección* (u *orden*) implícita. Como $a \leq b$ no necesariamente implica $b \leq a$, no debe suponer que $A \sim B$ implica $B \sim A$. (De hecho, esto *es* cierto, como se demostrará en el siguiente teorema, pero no se deduce inmediatamente de la definición.)
- La matriz P depende de A y B. No es única para un par dado de matrices A y B similares. Para ver esto, simplemente tome $A = B = I$, en cuyo caso $I \sim I$, pues $P^{-1}IP = I$ para *cualquier* matriz invertible P.

Ejemplo 4.22

Sea $A = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$. Entonces $A \sim B$, dado que

$$\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -2 & -1 \end{bmatrix}$$

Por tanto, $AP = PB$ con $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. (Note que no es necesario calcular P^{-1}. Vea el primer Comentario en la página anterior.)

Teorema 4.21 Sean A, B y C matrices de $n \times n$.

a. $A \sim A$
b. Si $A \sim B$, entonces $B \sim A$.
c. Si $A \sim B$ y $B \sim C$, entonces $A \sim C$.

Demostración (a) Esta propiedad surge del hecho de que $I^{-1}AI = A$.

(b) Si $A \sim B$, entonces $P^{-1}AP = B$ para alguna matriz invertible P. Como se hizo notar en el primer Comentario de la página anterior, esto es equivalente a $PBP^{-1} = A$. Al hacer $Q = P^{-1}$, se tiene $Q^{-1}BQ = (P^{-1})^{-1}BP^{-1} = PBP^{-1} = A$. Por tanto, por definición, $B \sim A$.

(c) En el ejercicio 30 se le pide demostrar la propiedad (c).

Comentario Cualquier relación que satisfaga las tres propiedades del Teorema 4.21 se llama ***relación de equivalencia***. Las relaciones de equivalencia surgen frecuentemente en matemáticas, y los objetos que se conectan vía una relación de este tipo por lo general comparten propiedades importantes. Está a punto de ver que esto también es cierto para las matrices semejantes.

Teorema 4.22 Sean A y B matrices de $n \times n$ con $A \sim B$. Entonces

a. $\det A = \det B$
b. A es invertible si y sólo si B es invertible.
c. A y B tienen el mismo rank.
d. A y B tienen el mismo polinomio característico.
e. A y B tienen los mismos eigenvalores.

Demostración Se demuestran (a) y (d) y las propiedades restantes se dejan como ejercicios. Si $A \sim B$, entonces $P^{-1}AP = B$ para alguna matriz invertible P.

(a) Al tomar determinantes de ambos lados, se tiene

$$\begin{aligned}\det B = \det(P^{-1}AP) &= (\det P^{-1})(\det A)(\det P)\\ &= \left(\frac{1}{\det P}\right)(\det A)(\det P) = \det A\end{aligned}$$

(d) El polinomio característico de B es

$$\begin{aligned}\det(B - \lambda I) &= \det(P^{-1}AP - \lambda I)\\ &= \det(P^{-1}AP - \lambda P^{-1}IP)\\ &= \det(P^{-1}AP - P^{-1}(\lambda I)P)\\ &= \det(P^{-1}(A - \lambda I)P) = \det(A - \lambda I)\end{aligned}$$

y el último paso procede como en (a). Por tanto, $\det(B - \lambda I) = \det(A - \lambda I)$; esto es, los polinomios característicos de B y A son los mismos.

Comentario Dos matrices pueden tener en común las propiedades (a) a (e) (y más) y aún así no ser semejantes. Por ejemplo, $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ tienen determinante 1 y rank 2, son invertibles y tienen polinomio característico $(1 - \lambda)^2$ y eigenvalores $\lambda_1 = \lambda_2 = 1$. Pero A no es semejante a B, pues $P^{-1}AP = P^{-1}IP = I \neq B$ para cualquier matriz invertible P.

El Teorema 4.22 es más útil para demostrar que dos matrices *no* son semejantes, pues A y B no pueden ser semejantes si falla alguna de las propiedades de la (a) a la (e).

Ejemplo 4.23

(a) $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ y $B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ no son semejantes, pues $\det A = -3$ pero $\det B = 3$.

(b) $A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ y $B = \begin{bmatrix} 1 & 1 \\ 3 & -1 \end{bmatrix}$ no son semejantes, pues el polinomio característico de A es $\lambda^2 - 3\lambda - 4$ mientras que el de B es $\lambda^2 - 4$. (Compruébelo.) Sin embargo, note que A y B tienen los mismos determinante y rank.

Diagonalización

La mejor situación posible es cuando una matriz cuadrada es semejante a una matriz diagonal. Como está a punto de ver, el que una matriz sea diagonalizable está estrechamente relacionado con los eigenvalores y eigenvectores de la matriz.

Definición Una matriz A de $n \times n$ es ***diagonalizable*** si existe una matriz diagonal D tal que A sea semejante a D; esto es, si existe una matriz invertible P de $n \times n$ tal que $P^{-1}AP = D$.

Ejemplo 4.24

$A = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$ es diagonalizable pues, si $P = \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix}$ y $D = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$, entonces $P^{-1}AP = D$, como puede comprobarse fácilmente. (En realidad, es más rápido comprobar el enunciado equivalente $AP = PD$, pues no requiere encontrar P^{-1}.)

El ejemplo 4.24 plantea la pregunta de dónde pueden provenir las matrices P y D. Observe que las entradas diagonales 4 y -1 de D son los eigenvalores de A, pues son las raíces de su polinomio característico, que se encontró en el ejemplo 4.23(b). El origen de la matriz P es menos obvio, pero, como está a punto de demostrarse, sus entradas se obtienen de los eigenvectores de A. El Teorema 4.23 precisa esta conexión.

Teorema 4.23 Sea A una matriz de $n \times n$. Entonces A es diagonalizable si y sólo si A tiene n eigenvectores linealmente independientes.

Más precisamente, existe una matriz invertible P y una matriz diagonal D tales que $P^{-1}AP = D$ si y sólo si las columnas de P son n eigenvectores linealmente independientes de A y las entradas de la diagonal de D son los eigenvalores de A correspondientes a los eigenvectores en P en el mismo orden.

Demostración Suponga primero que A es similar a la matriz diagonal D vía $P^{-1}AP = D$ o, de manera equivalente, $AP = PD$. Sean $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ las columnas de P y sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ las entradas de la diagonal de D. Entonces

$$A[\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n] = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \cdots \quad \mathbf{p}_n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{1}$$

o
$$[A\mathbf{p}_1 \quad A\mathbf{p}_2 \quad \cdots \quad A\mathbf{p}_n] = [\lambda_1\mathbf{p}_1 \quad \lambda_2\mathbf{p}_2 \quad \cdots \quad \lambda_n\mathbf{p}_n] \tag{2}$$

donde el lado derecho es justo la representación columna-renglón del producto PD. Al igualar las columnas se tiene

$$A\mathbf{p}_1 = \lambda_1\mathbf{p}_1, A\mathbf{p}_2 = \lambda_2\mathbf{p}_2, \ldots, A\mathbf{p}_n = \lambda_n\mathbf{p}_n$$

lo que demuestra que los vectores columna de P son eigenvectores de A cuyos eigenvalores correspondientes son las entradas diagonales de D en el mismo orden. Dado que P es invertible, sus columnas son linealmente independientes, por el teorema fundamental de las matrices invertibles.

Por el contrario, si A tiene n eigenvectores linealmente independientes $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ con eigenvalores correspondientes $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectivamente, entonces

$$A\mathbf{p}_1 = \lambda_1\mathbf{p}_1, A\mathbf{p}_2 = \lambda_2\mathbf{p}_2, \ldots, A\mathbf{p}_n = \lambda_n\mathbf{p}_n$$

Esto implica a la ecuación (2) anterior, que es equivalente a la ecuación (1). En consecuencia, si se toma P como la matriz de $n \times n$ con columnas $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$, entonces la ecuación (1) se convierte en $AP = PD$. Dado que las columnas de P son linealmente independientes, el teorema fundamental de las matrices invertibles implica que P es invertible, de modo que $P^{-1}AP = D$; esto es, A es diagonalizable.

Ejemplo 4.25

Si es posible, encuentre una matriz P que diagonalice

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$$

Solución Esta matriz se estudió en el ejemplo 4.18, donde se descubrió que tiene eigenvalores $\lambda_1 = \lambda_2 = 1$ y $\lambda_3 = 2$. Los eigenespacios tienen las siguientes bases:

$$\text{Para } \lambda_1 = \lambda_2 = 1, E_1 \text{ tiene base } \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

$$\text{Para } \lambda_3 = 2, E_2 \text{ tiene base } \begin{bmatrix} 1 \\ 2 \\ 4 \end{bmatrix}.$$

Dado que todos los demás eigenvectores son sólo múltiplos de uno de estos dos vectores base, no puede haber tres eigenvectores linealmente independientes. En consecuencia, por el Teorema 4.23, A no es diagonalizable.

Ejemplo 4.26

Si es posible, encuentre una matriz P que diagonalice

$$A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$$

Solución Esta es la matriz del ejemplo 4.19. Ahí se descubrió que los eigenvalores de A son $\lambda_1 = \lambda_2 = 0$ y $\lambda_3 = -2$, con las siguientes bases para los eigenespacios:

$$\text{Para } \lambda_1 = \lambda_2 = 0,\ E_0 \text{ tiene base } \mathbf{p}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \text{ y } \mathbf{p}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

$$\text{Para } \lambda_3 = -2,\ E_{-2} \text{ tiene base } \mathbf{p}_3 = \begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}.$$

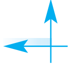
Se comprueba directamente que estos tres vectores son linealmente independientes. Por tanto, si se tiene

$$P = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3] = \begin{bmatrix} 0 & 1 & -1 \\ 1 & 0 & 3 \\ 0 & 1 & 1 \end{bmatrix}$$

entonces P es invertible. Más aún,

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{bmatrix} = D$$

como puede comprobarse fácilmente. (Si hace la comprobación a mano, es mucho más fácil comprobar la ecuación equivalente $AP = PD$.)



Comentarios

- Cuando existan suficientes eigenvectores, pueden colocarse en las columnas de P en cualquier orden. Sin embargo, los eigenvalores provendrán de la diagonal de D en el mismo orden que sus correspondientes eigenvectores en P. Por ejemplo, si hubiera elegido

$$P = [\mathbf{p}_1 \quad \mathbf{p}_3 \quad \mathbf{p}_2] = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 3 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

entonces habría encontrado

$$P^{-1}AP = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- En el ejemplo 4.26 se le pidió comprobar que los eigenvectores $\mathbf{p}_1, \mathbf{p}_2$, y $\mathbf{p}_3$ eran linealmente independientes. ¿Era necesario comprobar esto? Se sabía que $\{\mathbf{p}_1, \mathbf{p}_2\}$ era linealmente independiente, pues era una base para el eigenespacio E_0. También se sabía que los conjuntos $\{\mathbf{p}_1, \mathbf{p}_3\}$ y $\{\mathbf{p}_2, \mathbf{p}_3\}$ eran linealmente independientes, por el Teorema 4.20. Pero a partir de esta información *no* se podía concluir que $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ era linealmente independiente. Sin embargo, el siguiente teorema garantiza que la independencia lineal se conserva cuando se combinan las bases de diferentes eigenespacios.

Teorema 4.24 Sea A una matriz de $n \times n$ y sean $\lambda_1, \lambda_2, \ldots, \lambda_k$ distintos eigenvalores de A. Si $\mathcal{B}_i$ es una base para el eigenespacio E_{λ_i}, entonces $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \cdots \cup \mathcal{B}_k$ (es decir, la colección total de vectores base para todos los eigenespacios) es linealmente independiente.

Demostración Sea $\mathcal{B}_i = \{\mathbf{v}_{i1}, \mathbf{v}_{i2}, \ldots, \mathbf{v}_{in_i}\}$ para $i = 1, \ldots, k$. Tiene que demostrar que

$$\mathcal{B} = \{\mathbf{v}_{11}, \mathbf{v}_{12}, \ldots, \mathbf{v}_{1n_1}, \mathbf{v}_{21}, \mathbf{v}_{22}, \ldots, \mathbf{v}_{2n_2}, \ldots, \mathbf{v}_{k1}, \mathbf{v}_{k2}, \ldots, \mathbf{v}_{kn_k}\}$$

es linealmente independiente. Suponga que alguna combinación lineal no trivial de estos vectores es el vector cero, por decir,

$$(c_{11}\mathbf{v}_{11} + \cdots + c_{1n_1}\mathbf{v}_{1n_1}) + (c_{21}\mathbf{v}_{21} + \cdots + c_{2n_2}\mathbf{v}_{2n_2}) + \cdots + (c_{k1}\mathbf{v}_{k1} + \cdots + c_{kn_k}\mathbf{v}_{kn_k}) = \mathbf{0} \quad (3)$$

Al denotar las sumas entre paréntesis mediante $\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_k$, puede escribir la ecuación (3) como

$$\mathbf{x}_1 + \mathbf{x}_2 + \cdots + \mathbf{x}_k = \mathbf{0} \quad (4)$$



Ahora cada $\mathbf{x}_i$ está en E_{λ_i} (¿por qué?) y por tanto o es un eigenvector correspondiente a λ_i o es $\mathbf{0}$. Pero, puesto que los eigenvalores de λ_i son distintos, si *alguno* de los factores $\mathbf{x}_i$ es un eigenvector, son linealmente independientes, por el Teorema 4.20. Sin embargo, la ecuación (4) es una relación de *dependencia* lineal; esto es una contradicción. Se concluye que la ecuación (3) debe ser trivial; esto es: todos sus coeficientes son cero. Por tanto, $\mathcal{B}$ es linealmente independiente.

Existe un caso en el que la capacidad de diagonalización es automática: una matriz de $n \times n$ con n eigenvalores distintos.

Teorema 4.25 Si A es una matriz de $n \times n$ con n eigenvalores distintos, entonces A es diagonalizable.

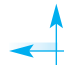
Demostración Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ eigenvectores correspondientes a los n eigenvalores distintos de A. (¿Por qué no podría haber *más* que n de tales eigenvectores?) Por el Teorema 4.20, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente independientes, de modo que, por el Teorema 4.23, A es diagonalizable.



Ejemplo 4.27 La matriz

$$A = \begin{bmatrix} 2 & -3 & 7 \\ 0 & 5 & 1 \\ 0 & 0 & -1 \end{bmatrix}$$

tiene eigenvalores $\lambda_1 = 2$, $\lambda_2 = 5$ y $\lambda_3 = -1$, por el Teorema 4.15. Dado que existen tres eigenvalores distintos para una matriz de 3×3, A es diagonalizable, por el Teorema 4.25. (Si en realidad requiere una matriz P tal que $P^{-1}AP$ sea diagonal, deberá calcular bases para los eigenespacios, como en el ejemplo 4.19 y el ejemplo 4.26 anteriores.)



El teorema final de esta sección es un importante resultado que caracteriza a las matrices diagonalizables en términos de las dos nociones de multiplicidad que se introdujeron después del ejemplo 4.18. Ofrece condiciones precisas con las cuales una matriz de $n \times n$ puede ser diagonalizable, aun cuando tenga menos de n eigenvalores, como en el ejemplo 4.26. Primero se prueba un lema que verifica si una matriz es o no diagonalizable.

Lema 4.26 Si A es una matriz de $n \times n$, entonces la multiplicidad geométrica de cada eigenvalor es menor o igual que su multiplicidad algebraica.

Demostración Suponga que λ_1 es un eigenvalor de A con multiplicidad geométrica p; esto es, $\dim E_{\lambda_1} = p$. De manera específica, sea E_{λ_1} que tiene base $\mathcal{B}_1 = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p\}$. Sea Q cualquier matriz invertible de $n \times n$ que tenga $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ como sus primeras columnas p, por ejemplo,

$$Q = [\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_p \quad \mathbf{v}_{p+1} \quad \cdots \quad \mathbf{v}_n]$$

o, como una matriz particionada,

$$Q = [U \mid V]$$

Sea

$$Q^{-1} = \begin{bmatrix} C \\ \hline D \end{bmatrix}$$

donde C es $p \times n$.

Dado que las columnas de U son eigenvectores que corresponden a λ_1, $AU = \lambda_1 U$. También se tiene

$$\left[\begin{array}{c|c} I_p & O \\ \hline O & I_{n-p} \end{array}\right] = I_n = Q^{-1}Q = \begin{bmatrix} C \\ \hline D \end{bmatrix}[U \mid V] = \left[\begin{array}{c|c} CU & CV \\ \hline DU & DV \end{array}\right]$$

de donde se obtiene $CU = I_p$, $CV = O$, $DU = O$ y $DV = I_{n-p}$. Por tanto,

$$Q^{-1}AQ = \begin{bmatrix} C \\ \hline D \end{bmatrix}A[U \mid V] = \left[\begin{array}{c|c} CAU & CAV \\ \hline DAU & DAV \end{array}\right] = \left[\begin{array}{c|c} \lambda_1 CU & CAV \\ \hline \lambda_1 DU & DAV \end{array}\right] = \left[\begin{array}{c|c} \lambda_1 I_p & CAV \\ \hline O & DAV \end{array}\right]$$

Por el ejercicio 69 de la sección 4.2, se tiene que

$$\det(Q^{-1}AQ - \lambda I) = (\lambda_1 - \lambda)^p \det(DAV - \lambda I) \tag{5}$$

Pero $\det(Q^{-1}AQ - \lambda I)$ es el polinomio característico de $Q^{-1}AQ$, que es igual que el polinomio característico de A, por el Teorema 4.22(d). Por tanto, la ecuación (5) implica que la multiplicidad algebraica de λ_1 es al menos p, su multiplicidad geométrica.

Teorema 4.27 **El teorema de diagonalización**

Sea A una matriz de $n \times n$ cuyos distintos eigenvalores son $\lambda_1, \lambda_2, \ldots, \lambda_k$. Los siguientes enunciados son equivalentes:

a. A es diagonalizable.
b. La unión $\mathcal{B}$ de las bases de los eigenespacios de A (como en el Teorema 4.24) contiene n vectores.
c. La multiplicidad algebraica de cada eigenvalor es igual a su multiplicidad geométrica.

Demostración (a) $\Rightarrow$ (b) Si A es diagonalizable, entonces tiene n eigenvectores linealmente independientes, por el Teorema 4.23. Si n_i de dichos eigenvectores corresponde al eigenvalor λ_i, entonces $\mathcal{B}_i$ contiene al menos n_i vectores. (Ya se sabe que dichos n_i vectores son linealmente independientes; la única cosa que puede evitar que sean una base de E_{λ_i} es que no puedan generarlo.) Por tanto, $\mathcal{B}$ contiene al menos n vectores. Pero, por el Teorema 4.24, $\mathcal{B}$ es un conjunto linealmente independiente en $\mathbb{R}^n$; por tanto, contiene exactamente n vectores.



(b) ⇒ (c) Sea $d_i = \dim E_{\lambda i}$ la multiplicidad geométrica de λ_i, y sea m_i la multiplicidad algebraica λ_i de m_i. Por el Lema 4.26, $d_i \leq m_i$ para $i = 1, \ldots, k$. Ahora suponga que se cumple la propiedad (b). Entonces también se tiene

$$n = d_1 + d_2 + \cdots + d_k \leq m_1 + m_2 + \cdots + m_k$$

Pero $m_1 + m_2 + \cdots + m_k = n$, pues la suma de las multiplicidades algebraicas de los eigenvalores de A es justo el grado del polinomio característico de A, a saber, n.

Se tiene que $d_1 + d_2 + \cdots + d_k = m_1 + m_2 + \cdots + m_k$, lo cual implica que

$$(m_1 - d_1) + (m_2 - d_2) + \cdots + (m_k - d_k) = 0 \qquad (6)$$

Al usar nuevamente el Lema 4.26, se sabe que $m_i - d_i \geq 0$ para $i = 1, \ldots, k$, de donde se puede deducir que cada sumando en la ecuación (6) es cero; esto es, $m_i = d_i$ para $i = 1, \ldots, k$.

(c) ⇒ (a) Si la multiplicidad algebraica m_i y la multiplicidad geométrica d_i son iguales para cada eigenvalor λ_i de A, entonces $\mathcal{B}$ tiene $d_1 + d_2 + \cdots + d_k = m_1 + m_2 + \cdots + m_k = n$ vectores que son linealmente independientes, por el Teorema 4.24. Por tanto, son n eigenvectores linealmente independientes de A, y A es diagonalizable, por el Teorema 4.23.

Ejemplo 4.28

(a) La matriz $A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}$ del ejemplo 4.18 tiene dos eigenvalores distintos: $\lambda_1 = \lambda_2 = 1$ y $\lambda_3 = 2$. Dado que el eigenvalor $\lambda_1 = \lambda_2 = 1$ tiene multiplicidad algebraica 2 pero multiplicidad geométrica 1, A no es diagonalizable, por el teorema de diagonalización. (Vea también el ejemplo 4.25.)

(b) La matriz $A = \begin{bmatrix} -1 & 0 & 1 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{bmatrix}$ del ejemplo 4.19 también tiene dos eigenvalores distintos, $\lambda_1 = \lambda_2 = 0$ y $\lambda_3 = -2$. El eigenvalor 0 tiene multiplicidades algebraica y geométrica 2, y el eigenvalor -2 tiene multiplicidades algebraica y geométrica 1. Por tanto, esta matriz *es* diagonalizable, por el teorema de diagonalización. (Esto concuerda con los hallazgos del ejemplo 4.26.)

Esta sección concluye con una aplicación de la diagonalización al cálculo de las potencias de una matriz.

Ejemplo 4.29

Calcule A^{10} si $A = \begin{bmatrix} 0 & 1 \\ 2 & 1 \end{bmatrix}$.

Solución En el ejemplo 4.21 se descubrió que esta matriz tiene eigenvalores $\lambda_1 = -1$ y $\lambda_2 = 2$, con eigenvectores correspondientes $\mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Se tiene (de alguno de los teoremas de esta sección) que A es diagonalizable y $P^{-1}AP = D$, donde

$$P = [\mathbf{v}_1 \quad \mathbf{v}_2] = \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix} \qquad \text{y} \qquad D = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$$

Al despejar para A, se tiene $A = PDP^{-1}$, que facilita encontrar potencias de A. Calcule

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD(P^{-1}P)DP^{-1} = PDIDP^{-1} = PD^2P^{-1}$$



y, por lo general, $A^n = PD^nP^{-1}$ para todo $n \geq 1$. (Debe verificar esto por inducción. Observe que este hecho será verdadero para *cualquier* matriz diagonalizable, no sólo para la de este ejemplo.)

Dado que

$$D^n = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}^n = \begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix}$$

se tiene

$$\begin{aligned} A^n = PD^nP^{-1} &= \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} (-1)^n & 0 \\ 0 & 2^n \end{bmatrix}\begin{bmatrix} \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{bmatrix} \\ &= \begin{bmatrix} \dfrac{2(-1)^n + 2^n}{3} & \dfrac{(-1)^{n+1} + 2^n}{3} \\ \dfrac{2(-1)^{n+1} + 2^{n+1}}{3} & \dfrac{(-1)^{n+2} + 2^{n+1}}{3} \end{bmatrix} \end{aligned}$$

Puesto que sólo se pidió A^{10}, esto es más de lo que necesita. Pero ahora puede simplemente establecer $n = 10$ para encontrar

$$A^{10} = \begin{bmatrix} \dfrac{2(-1)^{10} + 2^{10}}{3} & \dfrac{(-1)^{11} + 2^{10}}{3} \\ \dfrac{2(-1)^{11} + 2^{11}}{3} & \dfrac{(-1)^{12} + 2^{11}}{3} \end{bmatrix} = \begin{bmatrix} 342 & 341 \\ 682 & 683 \end{bmatrix}$$



Ejercicios 4.4

En los ejercicios 1-4, demuestre que A y B no son matrices semejantes.

1. $A = \begin{bmatrix} 4 & 1 \\ 3 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, B = \begin{bmatrix} 2 & -5 \\ -2 & 4 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & 1 & 4 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 4 & 0 \\ 2 & 3 & 4 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 1 & 4 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 2 & 2 \\ 1 & 0 & 3 \end{bmatrix}$

En los ejercicios 5-7, una diagonalización de la matriz A se da en la forma $P^{-1}AP = D$. Mencione los eigenvalores de A y las bases para los eigenespacios correspondientes.

5. $\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 5 & -1 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 3 \end{bmatrix}$

6. $\begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{1}{3} \\ 0 & -1 & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 1 & 2 \end{bmatrix}$

$= \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

7. $\begin{bmatrix} \frac{1}{8} & \frac{1}{8} & \frac{1}{8} \\ -\frac{1}{4} & \frac{3}{4} & -\frac{1}{4} \\ \frac{5}{8} & -\frac{3}{8} & -\frac{3}{8} \end{bmatrix} \begin{bmatrix} 1 & 3 & 3 \\ 2 & 0 & 2 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 1 \\ 2 & 1 & 0 \\ 3 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

En los ejercicios 8-15, determine si A es diagonalizable y, si lo es, encuentre una matriz invertible P y una matriz diagonal D tales que $P^{-1}AP = D$.

8. $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

9. $A = \begin{bmatrix} -3 & 4 \\ -1 & 1 \end{bmatrix}$

10. $A = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{bmatrix}$

11. $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

12. $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 3 \\ 3 & 0 & 1 \end{bmatrix}$

13. $A = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

14. $A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 2 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

15. $A = \begin{bmatrix} 2 & 0 & 0 & 4 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$

En los ejercicios 16-23, use el método del ejemplo 4.29 para calcular la potencia indicada de la matriz.

16. $\begin{bmatrix} -5 & 8 \\ -4 & 7 \end{bmatrix}^5$

17. $\begin{bmatrix} -1 & 6 \\ 1 & 0 \end{bmatrix}^{10}$

18. $\begin{bmatrix} 0 & -3 \\ -1 & 2 \end{bmatrix}^{-5}$

19. $\begin{bmatrix} 0 & 3 \\ 1 & 2 \end{bmatrix}^k$

20. $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}^8$

21. $\begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}^{2002}$

22. $\begin{bmatrix} 3 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 3 \end{bmatrix}^k$

23. $\begin{bmatrix} 1 & 1 & 0 \\ 2 & -2 & 2 \\ 0 & 1 & 1 \end{bmatrix}^k$

En los ejercicios 24-29, encuentre todos los valores (reales) de k para los cuales A sea diagonalizable.

24. $A = \begin{bmatrix} 1 & 1 \\ 0 & k \end{bmatrix}$

25. $A = \begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$

26. $A = \begin{bmatrix} k & 1 \\ 1 & 0 \end{bmatrix}$

27. $A = \begin{bmatrix} 1 & 0 & k \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

28. $A = \begin{bmatrix} 1 & k & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

29. $A = \begin{bmatrix} 1 & 1 & k \\ 1 & 1 & k \\ 1 & 1 & k \end{bmatrix}$

30. Demuestre el Teorema 4.21(c).

31. Demuestre el Teorema 4.22(b).

32. Demuestre el Teorema 4.22(c).

33. Demuestre el Teorema 4.22(e).

34. Si A y B son matrices invertibles, demuestre que AB y BA son semejantes.

35. Demuestre que si A y B son matrices semejantes, entonces $\operatorname{tr}(A) = \operatorname{tr}(B)$. [*Sugerencia:* encuentre una forma de usar el ejercicio 45 de la sección 3.2.]

En general, es difícil demostrar que dos matrices son semejantes. Sin embargo, si dos matrices semejantes son diagonalizables, la tarea se vuelve más sencilla. En los ejercicios 36-39, demuestre que A y B son semejantes al demostrar que son semejantes a la misma matriz diagonal. Luego encuentre una matriz invertible P tal que $P^{-1}AP = B$.

36. $A = \begin{bmatrix} 3 & 1 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

37. $A = \begin{bmatrix} 5 & -3 \\ 4 & -2 \end{bmatrix}, B = \begin{bmatrix} -1 & 1 \\ -6 & 4 \end{bmatrix}$

38. $A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 3 & 2 & -5 \\ 1 & 2 & -1 \\ 2 & 2 & -4 \end{bmatrix}$

39. $A = \begin{bmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \\ 2 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} -3 & -2 & 0 \\ 6 & 5 & 0 \\ 4 & 4 & -1 \end{bmatrix}$

40. Demuestre que si A es semejante a B, entonces A^T es semejante a B^T.

41. Demuestre que si A es diagonalizable, también lo es A^T.

42. Sea A una matriz invertible. Demuestre que, si A es diagonalizable, también lo es A^{-1}.

43. Demuestre que si A es una matriz diagonalizable con un solo eigenvalor λ, entonces A es de la forma $A = \lambda I$. (A tal matriz se le llama ***matriz escalar***.)

44. Sean A y B matrices de $n \times n$, cada una con n eigenvalores distintos. Pruebe que A y B tienen los mismos eigenvectores si y sólo si $AB = BA$.

45. Sean A y B matrices semejantes. Pruebe que las multiplicidades algebraicas de los eigenvalores de A y B son iguales.

46. Sean A y B matrices semejantes. Pruebe que las multiplicidades geométricas de los eigenvalores de A y B son iguales. [*Sugerencia:* demuestre que si $B = P^{-1}AP$, entonces todo eigenvector de B es de la forma $P^{-1}\mathbf{v}$ para algún eigenvector $\mathbf{v}$ de A.]

47. Demuestre que si A es una matriz diagonalizable tal que todo eigenvalor de A es 0 o 1, entonces A es idempotente (esto es: $A^2 = A$).

48. Sea A una matriz nilpotente (esto es: $A^m = O$ para algún $m > 1$). Pruebe que si A es diagonalizable, entonces A debe ser la matriz ero.

49. Suponga que A es una matriz de 6×6 con polinomio característico $c_A(\lambda) = (1 + \lambda)(1 - \lambda)^2(2 - \lambda)^3$.

(a) Demuestre que no es posible encontrar tres vectores linealmente independientes $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ en $\mathbb{R}^6$ tales que $A\mathbf{v}_1 = \mathbf{v}_1, A\mathbf{v}_2 = \mathbf{v}_2$ y $A\mathbf{v}_3 = \mathbf{v}_3$.

(b) Si A es diagonalizable, ¿cuáles son las dimensiones de los eigenespacios E_{-1}, E_1 y E_2?

50. Sea $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

(a) Demuestre que A es diagonalizable si $(a - d)^2 + 4bc > 0$ y no es diagonalizable si $(a - d)^2 + 4bc < 0$.

(b) Encuentre dos ejemplos para demostrar que si $(a - d)^2 + 4bc = 0$, entonces A puede o no ser diagonalizable.

4.5 Métodos iterativos para calcular eigenvalores

En este punto, el único método que se tiene para calcular los eigenvalores de una matriz es resolver la ecuación característica. Sin embargo, existen muchos problemas con este método que resultan poco prácticos excepto para los ejemplos pequeños. El primer problema es que depende del cálculo de un determinante, que es un proceso que consume mucho tiempo para matrices grandes. El segundo problema es que la ecuación característica es polinomial, y no hay fórmulas para resolver ecuaciones polinomiales de grado mayor que 4 (los polinomios de grados 2, 3 y 4 pueden resolverse usando la fórmula cuadrática y sus análogas). Por tanto, uno se ve forzado a *aproximar* eigenvalores en la mayoría de los problemas prácticos. Por desgracia, los métodos para aproximar las raíces de un polinomio son muy sensibles a errores de redondeo y por tanto no son confiables.

En lugar de ello, se pasa por alto el polinomio característico y se toma un enfoque diferente, al aproximar primero un eigenvector y luego usar este eigenvector para encontrar el eigenvalor correspondiente. En esta sección se explorarán algunas variaciones de tal método, que se basa en una simple técnica iterativa.

En 1824, el matemático noruego Niels Henrik Abel (1802-1829) demostró que una ecuación polinomial general de quinto grado (quíntica) no se puede *resolver mediante radicales*; esto es: no hay una fórmula para sus raíces en términos de sus coeficientes que use solamente las operaciones de suma, resta, multiplicación, división y sacar n-ésimas raíces. En un ensayo escrito en 1830 y publicado de manera póstuma en 1846, el matemático francés Evariste Galois (1811-1832) ofreció una teoría más completa que establecía condiciones bajo las cuales una ecuación polinominal arbitraria puede resolverse mediante radicales. El trabajo de Galois fue útil para el establecimiento de la rama del álgebra llamada *teoría de grupos*; su enfoque de las ecuaciones polinomiales ahora se conoce como *teoría de Galois*.

El método de potencia

El método de potencia se aplica a una matriz de $n \times n$ que tiene un ***eigenvalor dominante*** λ_1; esto es, un eigenvalor que es mayor en valor absoluto que todos los otros eigenvalores. Por ejemplo, si una matriz tiene eigenvalores $-4, -3, 1$ y 3, entonces -4 es el eigenvalor dominante, pues $4 = |-4| > |-3| \geq |3| \geq |1|$. Por otra parte, una matriz con eigenvalores $-4, -3, 3$ y 4 no tiene eigenvalor dominante.

El método de potencia avanza de manera iterativa para producir una secuencia de escalares que converja en λ_1 y una secuencia de vectores que converja en el correspondiente eigenvector $\mathbf{v}_1$, el ***eigenvector dominante***. Por simplicidad, se supondrá que la matriz A es diagonalizable. El siguiente teorema es la base para el método de potencia.

Teorema 4.28

Sea A una matriz diagonalizable de $n \times n$ con eigenvalor dominante λ_1. Entonces existe un vector distinto de cero $\mathbf{x}_0$ tal que la sucesión de vectores $\mathbf{x}_k$ definida por

$$\mathbf{x}_1 = A\mathbf{x}_0, \mathbf{x}_2 = A\mathbf{x}_1, \mathbf{x}_3 = A\mathbf{x}_2, \ldots, \mathbf{x}_k = A\mathbf{x}_{k-1}, \ldots$$

se aproxima a un eigenvector dominante de A.

Demostración Puede suponer que los eigenvalores de A se han marcado de modo que

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$$

Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ los correspondientes eigenvectores. Dado que $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ son linealmente independientes (¿por qué?), forman una base para $\mathbb{R}^n$. En consecuencia, puede escribir $\mathbf{x}_0$ como una combinación lineal de dichos eigenvectores, por ejemplo,

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

Ahora, $\mathbf{x}_1 = A\mathbf{x}_0$, $\mathbf{x}_2 = A\mathbf{x}_1 = A(A\mathbf{x}_0) = A^2\mathbf{x}_0$, $\mathbf{x}_3 = A\mathbf{x}_2 = A(A^2\mathbf{x}_0) = A^3\mathbf{x}_0$ y, por lo general,

$$\mathbf{x}_k = A^k\mathbf{x}_0 \quad \text{para } k \geq 1$$

Como se vio en el ejemplo 4.21,

$$\begin{aligned} A^k\mathbf{x}_0 &= c_1\lambda_1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \cdots + c_n\lambda_n^k\mathbf{v}_n \\ &= \lambda_1^k\left(c_1\mathbf{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^k\mathbf{v}_2 + \cdots + c_n\left(\frac{\lambda_n}{\lambda_1}\right)^k\mathbf{v}_n\right) \end{aligned} \tag{1}$$

donde se usó el hecho de que $\lambda_1 \neq 0$

El hecho de que λ_1 sea el eigenvalor dominante significa que cada una de las fracciones $\lambda_2/\lambda_1, \lambda_3/\lambda_1, \ldots, \lambda_n/\lambda_1$, es menor que 1 en valor absoluto. Por tanto,

$$\left(\frac{\lambda_2}{\lambda_1}\right)^k, \left(\frac{\lambda_3}{\lambda_1}\right)^k, \ldots, \left(\frac{\lambda_n}{\lambda_1}\right)^k$$

todos tienden a cero conforme $k \to \infty$. Se tiene que

$$\mathbf{x}_k = A^k\mathbf{x}_0 \to \lambda_1^k c_1\mathbf{v}_1 \quad \text{a medida que } k \to \infty \tag{2}$$

Ahora, dado que $\lambda_1 \neq 0$ y $\mathbf{v}_1 \neq \mathbf{0}$, $\mathbf{x}_k$ tiende a un múltiplo *distinto de cero* de $\mathbf{v}_1$ (esto es, un eigenvector que corresponde a λ_1) *siempre que* $c_1 \neq 0$. (Esta es la condición requerida acerca del vector inicial $\mathbf{x}_0$: debe tener un componente distinto de cero c_1 en la dirección del eigenvector dominante $\mathbf{v}_1$.)

Ejemplo 4.30

Aproxime el eigenvector dominante de $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ usando el método del Teorema 4.28.

Solución Tome $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ como el vector inicial. Entonces

$$\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Siga de esta forma para obtener los valores de $\mathbf{x}_k$ en la tabla 4.1.

Tabla 4.1

k	0	1	2	3	4	5	6	7	8
$\mathbf{x}_k$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\2\end{bmatrix}$	$\begin{bmatrix}3\\2\end{bmatrix}$	$\begin{bmatrix}5\\6\end{bmatrix}$	$\begin{bmatrix}11\\10\end{bmatrix}$	$\begin{bmatrix}21\\22\end{bmatrix}$	$\begin{bmatrix}43\\42\end{bmatrix}$	$\begin{bmatrix}85\\86\end{bmatrix}$	$\begin{bmatrix}171\\170\end{bmatrix}$
r_k	–	0.50	1.50	0.83	1.10	0.95	1.02	0.99	1.01
l_k	–	1.00	3.00	1.67	2.20	1.91	2.05	1.98	2.01

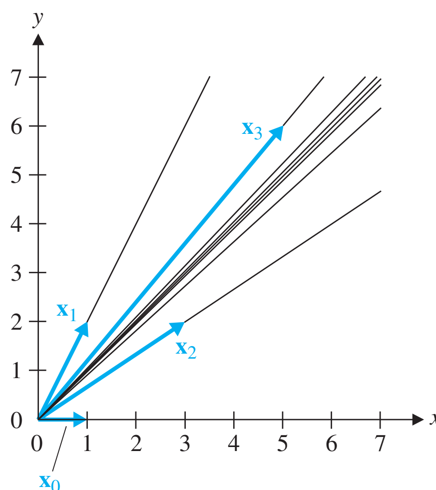


Figura 4.14

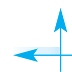


La figura 4.14 muestra lo que ocurre geométricamente. Se sabe que el eigenespacio para el eigenvector dominante tendrá dimensión 1. (¿Por qué? Vea el ejercicio 46.) Por tanto, es una recta que pasa por el origen en $\mathbb{R}^2$. Se muestran las primeras iteraciones $\mathbf{x}_k$ junto con las direcciones que determinan. Parece como si las iteraciones convergieran sobre la línea cuyo vector director es $\begin{bmatrix}1\\1\end{bmatrix}$. Para confirmar que este es el eigenvector dominante que se busca, sólo es necesario observar que la razón r_k del primero al segundo componente de $\mathbf{x}_k$ se acerca mucho a 1 conforme k aumenta. La segunda línea en el cuerpo de la tabla 4.1 da dichos valores y puede ver claramente que r_k de hecho tiende a 1. Se deduce que un eigenvector dominante de A es $\begin{bmatrix}1\\1\end{bmatrix}$.

Una vez encontrado un eigenvector dominante, ¿cómo puede encontrar el correspondiente eigenvalor dominante? Un método es observar que si un $\mathbf{x}_k$ tiende a un eigenvector dominante de A para el eigenvalor dominante λ_1, entonces

$$\mathbf{x}_{k+1} = A\mathbf{x}_k \approx \lambda_1\mathbf{x}_k$$

Se tiene que la razón l_k entre el primer componente de $\mathbf{x}_{k+1}$ y el de $\mathbf{x}_k$ tenderá a λ_1 conforme k aumente. La tabla 4.1 presenta los valores de l_k y puede ver que tienden a 2, que es el eigenvalor dominante.



Hay un inconveniente con el método del ejemplo 4.30: los componentes de las iteraciones $\mathbf{x}_k$ se vuelven muy grandes muy rápidamente y pueden causar errores significativos de redondeo. Para evitar este inconveniente, puede multiplicar cada iteración por algún escalar que reduzca la magnitud de sus componentes. Dado que los múltiplos escalares de las iteraciones $\|\mathbf{x}_k\|$ aún convergen a un eigenvector dominante, este enfoque es aceptable. Existen varias formas de lograr esto. Un método más sencillo, y el que se usará aquí, es dividir cada $\mathbf{x}_k$ entre el componente con el máximo valor absoluto, de modo que el componente más grande ahora sea 1. Este método se llama ***escalamiento***. Por tanto, si m_k denota el componente de $\mathbf{x}_k$ con el máximo valor absoluto, se sustituirá $\mathbf{x}_k$ con $\mathbf{y}_k = (1/m_k)\mathbf{x}_k$.

Este enfoque se ilustra con los cálculos del ejemplo 4.30. Para $\mathbf{x}_0$ no hay nada que hacer, pues $m_0 = 1$. En consecuencia,

$$\mathbf{y}_0 = \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Luego calcule $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ como antes, pero ahora escale con $m_1 = 2$ para obtener

$$\mathbf{y}_1 = \left(\frac{1}{2}\right)\mathbf{x}_1 = \left(\frac{1}{2}\right)\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$$

Ahora cambian los cálculos. Tome

$$\mathbf{x}_2 = A\mathbf{y}_1 = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$$

y escale para obtener

$$\mathbf{y}_2 = \left(\frac{1}{1.5}\right)\begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.67 \end{bmatrix}$$

Los siguientes cálculos se resumen en la tabla 4.2.

Ahora puede ver claramente que la sucesión de vectores $\mathbf{y}_k$ converge a $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, un eigenvector dominante. Más aún, la sucesión de escalares m_k converge en el correspondiente eigenvalor dominante $\lambda_1 = 2$.

Tabla 4.2

k	0	1	2	3	4	5	6	7	8
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1.67 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 1.83 \\ 1.67 \end{bmatrix}$	$\begin{bmatrix} 1.91 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 1.95 \\ 1.91 \end{bmatrix}$	$\begin{bmatrix} 1.98 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 1.99 \\ 1.98 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.67 \end{bmatrix}$	$\begin{bmatrix} 0.83 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.91 \end{bmatrix}$	$\begin{bmatrix} 0.95 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.98 \end{bmatrix}$	$\begin{bmatrix} 0.99 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.99 \end{bmatrix}$
m_k	1	2	1.5	2	1.83	2	1.95	2	1.99

A continuación se resume este método, llamado ***método de potencia***.

El método de potencia

Sea A una matriz de $n \times n$ diagonalizable con un correspondiente eigenvalor dominante λ_1.

1. Sea $\mathbf{x}_0 = \mathbf{y}_0$ cualquier vector inicial en $\mathbb{R}^n$ cuyo componente más grande sea 1.
2. Repita los siguientes pasos para $k = 1, 2, \ldots$:
   (a) Calcule $\mathbf{x}_k = A\mathbf{y}_{k-1}$.
   (b) Sea m_k el componente de $\mathbf{x}_k$ con el valor absoluto más grande.
   (c) Establezca $\mathbf{y}_k = (1/m_k)\mathbf{x}_k$.

Para la mayoría de las elecciones de $\mathbf{x}_0$, m_k converge en el eigenvalor dominante λ_1 y $\mathbf{y}_k$ converge en un eigenvector dominante.

Ejemplo 4.31

Use el método de potencia para aproximar el eigenvalor dominante y un eigenvector dominante de

$$A = \begin{bmatrix} 0 & 5 & -6 \\ -4 & 12 & -12 \\ -2 & -2 & 10 \end{bmatrix}$$

Solución Al tomar como vector inicial

$$\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

calcule las entradas de la tabla 4.3.

Puede ver que los vectores $\mathbf{y}_k$ se aproximan a $\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$ y los escalares m_k se aproximan a 16. Esto sugiere que son, respectivamente, un eigenvector dominante y el eigenvalor dominante de A.

Comentarios

- Si el vector inicial $\mathbf{x}_0$ tiene un componente cero en la dirección del eigenvector dominante $\mathbf{v}_1$ (esto es, si $c = 0$ en la demostración del Teorema 4.28), entonces el método de potencia no convergerá en un eigenvector dominante. Sin embargo, es muy probable

Tabla 4.3

k	0	1	2	3	4	5	6	7
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -1 \\ -4 \\ 6 \end{bmatrix}$	$\begin{bmatrix} -9.33 \\ -19.33 \\ 11.67 \end{bmatrix}$	$\begin{bmatrix} 8.62 \\ 17.31 \\ -9.00 \end{bmatrix}$	$\begin{bmatrix} 8.12 \\ 16.25 \\ -8.20 \end{bmatrix}$	$\begin{bmatrix} 8.03 \\ 16.05 \\ -8.04 \end{bmatrix}$	$\begin{bmatrix} 8.01 \\ 16.01 \\ -8.01 \end{bmatrix}$	$\begin{bmatrix} 8.00 \\ 16.00 \\ -8.00 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.17 \\ -0.67 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.48 \\ 1 \\ -0.60 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.52 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$	$\begin{bmatrix} 0.50 \\ 1 \\ -0.50 \end{bmatrix}$
m_k	1	6	−19.33	17.31	16.25	16.05	16.01	16.00

que durante el cálculo de las iteraciones posteriores, en algún punto el error de redondeo producirá un $\mathbf{x}_k$ con un componente distinto de cero en la dirección de $\mathbf{v}_1$. El método de potencia comenzará entonces a converger a un múltiplo de $\mathbf{v}_1$. (¡Este es un ejemplo en el que los errores de redondeo en realidad ayudan!)

- El método de potencia todavía funciona cuando hay un eigenvalor dominante *repetido* o incluso cuando la matriz no es diagonalizable, bajo ciertas condiciones. Puede encontrar detalles en la mayoría de los libros modernos acerca de análisis numérico. (Vea los ejercicios 21-24.)
- Para algunas matrices el método de potencia converge rápidamente a un eigenvector dominante, mientras para otros la convergencia puede ser muy lenta. Un vistazo cuidadoso a la demostración del Teorema 4.28 revela por qué. Dado que $|\lambda_2/\lambda_1| \geq |\lambda_3/\lambda_1| \geq \cdots \geq |\lambda_n/\lambda_1|$, si $|\lambda_2/\lambda_1|$ está cerca de cero, entonces $(\lambda_2/\lambda_1)^k, \ldots, (\lambda_n/\lambda_1)^k$ tenderán todos a cero rápidamente. La ecuación (2) muestra entonces que $\mathbf{x}_k = A^k\mathbf{x}_0$ también tenderá rápidamente a $\lambda_1^k c_1 \mathbf{v}_1$.

John William Strutt (1842-1919), barón Rayleigh, fue un físico británico que realizó grandes aportaciones a los campos de la acústica y óptica. En 1871 proporcionó la primera explicación correcta de por qué el cielo es azul, y en 1895 descubrió el gas inerte argón, por cuyo descubrimiento recibió el Premio Nobel en 1904. Rayleigh fue presidente de la Royal Society, de 1905 a 1908, y se convirtió en rector de la Universidad de Cambridge en 1908. Usó los cocientes de Rayleigh en un artículo de 1873 acerca de sistemas vibratorios y más tarde en su libro *The Theory of Sound*.

Como ilustración, considere el ejemplo 4.31. Los eigenvalores son 16, 4 y 2, de modo que $\lambda_2/\lambda_1 = 4/16 = 0.25$. Dado que $0.25^7 \approx 0.00006$, hacia la séptima iteración se estaría cerca de una precisión de cuatro lugares decimales. Esto es exactamente lo que se vio.

- Hay una forma alternativa de estimar el eigenvalor dominante λ_1 de una matriz A en conjunción con el método de potencia. Primero, observe que si $A\mathbf{x} = \lambda_1\mathbf{x}$, entonces

$$\frac{(A\mathbf{x}) \cdot \mathbf{x}}{\mathbf{x} \cdot \mathbf{x}} = \frac{(\lambda_1\mathbf{x}) \cdot \mathbf{x}}{\mathbf{x} \cdot \mathbf{x}} = \frac{\lambda_1(\mathbf{x} \cdot \mathbf{x})}{\mathbf{x} \cdot \mathbf{x}} = \lambda_1$$

La expresión $R(\mathbf{x}) = ((A\mathbf{x}) \cdot \mathbf{x})/(\mathbf{x} \cdot \mathbf{x})$ se llama ***cociente de Rayleigh***. Conforme se calculan las iteraciones $\mathbf{x}_k$, los sucesivos cocientes de Rayleigh $R(\mathbf{x}_k)$ deben tender a λ_1. De hecho, para matrices simétricas, el método de cocientes de Rayleigh es aproximadamente el doble de rápido que el método de factor de escalamiento. (Vea los ejercicios 17-20.)

El método de potencia ajustado y el método de potencia inverso

El método de potencia puede ayudarlo a aproximar el eigenvalor *dominante* de una matriz, ¿pero qué debe hacerse si se quieren los otros eigenvalores? Por fortuna, existen muchas variaciones del método de potencia que pueden aplicarse.

El ***método de potencia ajustado*** usa la observación de que si λ es un eigenvalor de A, entonces $\lambda - \alpha$ es un eigenvalor de $A - \alpha I$ para cualquier escalar α (ejercicio 22 de la sección 4.3). Por tanto, si λ_1 es el eigenvalor dominante de A, los eigenvalores de $A - \lambda_1 I$ serán $0, \lambda_2 - \lambda_1, \lambda_3 - \lambda_1, \ldots, \lambda_n - \lambda_1$. Entonces puede aplicar el método de potencia para calcular $\lambda_2 - \lambda_1$, y a partir de este valor puede encontrar λ_2. La repetición de este proceso le permitirá calcular todos los eigenvalores.

Ejemplo 4.32 Use el método de potencia ajustado para calcular el segundo eigenvalor de la matriz $A = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}$ del ejemplo 4.30.

Solución En el ejemplo 4.30 se encontró que $\lambda_1 = 2$. Para encontrar λ_2, aplique el método de potencia a

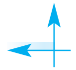
$$A - 2I = \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}$$

Se tomará $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, pero otras elecciones también funcionarán. Los cálculos se resumen en la tabla 4.4.

Tabla 4.4

k	0	1	2	3	4
$\mathbf{x}_k$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}-1\\2\end{bmatrix}$	$\begin{bmatrix}1.5\\-3\end{bmatrix}$	$\begin{bmatrix}1.5\\-3\end{bmatrix}$	$\begin{bmatrix}1.5\\-3\end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}-0.5\\1\end{bmatrix}$	$\begin{bmatrix}-0.5\\1\end{bmatrix}$	$\begin{bmatrix}-0.5\\1\end{bmatrix}$	$\begin{bmatrix}-0.5\\1\end{bmatrix}$
m_k	1	2	-3	-3	-3

La elección de $\mathbf{x}_0$ produjo el eigenvalor -3 después de sólo dos iteraciones. Por tanto, $\lambda_2 - \lambda_1 = -3$, de modo que $\lambda_2 = \lambda_1 - 3 = 2 - 3 = -1$ es el segundo eigenvalor de A.



Recuerde de la propiedad (b) del Teorema 4.18 que, si A es invertible con eigenvalor λ, entonces A^{-1} tiene eigenvalor $1/\lambda$. En consecuencia, si aplica el método de potencia a A^{-1}, su eigenvalor dominante será el *recíproco del más pequeño* (en magnitud) eigenvalor de A. Para usar este ***método de potencia inverso***, se siguen los mismos pasos que en el método de potencia, excepto que en el paso 2(a), calcule $\mathbf{x}_k = A^{-1}\mathbf{y}_{k-1}$. (En la práctica en realidad no se calcula A^{-1} explícitamente; en vez de ello, se resuelve la ecuación equivalente $A\mathbf{x}_k = \mathbf{y}_{k-1}$ para $\mathbf{x}_k$ usando eliminación gaussiana. Esto resulta ser más rápido.)

Ejemplo 4.33

Use el método de potencia inverso para calcular el segundo eigenvalor de la matriz $A = \begin{bmatrix}1 & 1\\2 & 0\end{bmatrix}$ del ejemplo 4.30.

Solución Comience, como en el ejemplo 4.30, con $\mathbf{x}_0 = \mathbf{y}_0 = \begin{bmatrix}1\\0\end{bmatrix}$. Para resolver $A\mathbf{x}_1 = \mathbf{y}_0$, use reducción por renglones:

$$[A \mid \mathbf{y}_0] = \left[\begin{array}{cc|c}1 & 1 & 1\\2 & 0 & 0\end{array}\right] \longrightarrow \left[\begin{array}{cc|c}1 & 0 & 0\\0 & 1 & 1\end{array}\right]$$

Por tanto, $\mathbf{x}_1 = \begin{bmatrix}0\\1\end{bmatrix}$, de modo que $\mathbf{y}_1 = \begin{bmatrix}0\\1\end{bmatrix}$. Entonces se obtiene $\mathbf{x}_2$ a partir de $A\mathbf{x}_2 = \mathbf{y}_1$:

$$[A \mid \mathbf{y}_1] = \left[\begin{array}{cc|c}1 & 1 & 0\\2 & 0 & 1\end{array}\right] \longrightarrow \left[\begin{array}{cc|c}1 & 0 & 0.5\\0 & 1 & -0.5\end{array}\right]$$

Por tanto, $\mathbf{x}_2 = \begin{bmatrix}0.5\\-0.5\end{bmatrix}$ y, por escalamiento, se tiene $\mathbf{y}_2 = \begin{bmatrix}1\\-1\end{bmatrix}$. Al continuar, se obtienen los valores que se muestran en la tabla 4.5, donde los valores m_k convergen en -1. Por tanto, el eigenvalor más pequeño de A es el recíproco de -1 (que también es -1). Esto concuerda con el hallazgo previo en el ejemplo 4.32.

Tabla 4.5

k	0	1	2	3	4	5	6	7	8	9
$\mathbf{x}_k$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}0.5\\-0.5\end{bmatrix}$	$\begin{bmatrix}-0.5\\1.5\end{bmatrix}$	$\begin{bmatrix}0.5\\-0.83\end{bmatrix}$	$\begin{bmatrix}0.5\\-1.1\end{bmatrix}$	$\begin{bmatrix}0.5\\-0.95\end{bmatrix}$	$\begin{bmatrix}0.5\\-1.02\end{bmatrix}$	$\begin{bmatrix}0.5\\-0.99\end{bmatrix}$	$\begin{bmatrix}0.5\\-1.01\end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix}0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\-1\end{bmatrix}$	$\begin{bmatrix}-0.33\\1\end{bmatrix}$	$\begin{bmatrix}-0.6\\1\end{bmatrix}$	$\begin{bmatrix}-0.45\\1\end{bmatrix}$	$\begin{bmatrix}-0.52\\1\end{bmatrix}$	$\begin{bmatrix}-0.49\\1\end{bmatrix}$	$\begin{bmatrix}-0.51\\1\end{bmatrix}$	$\begin{bmatrix}-0.50\\1\end{bmatrix}$
m_k	1	1	0.5	1.5	−0.83	−1.1	−0.95	−1.02	−0.99	−1.01

El método de potencia inverso ajustado

La más versátil de las variantes del método de potencia es una que combina los dos recién mencionados. Puede usarse para encontrar una aproximación para *cualquier* eigenvalor, siempre que se tenga una aproximación cercana a dicho eigenvalor. En otras palabras, si se proporciona un escalar α, el ***método de potencia inverso ajustado*** encontrará el eigenvalor λ de A que esté más cercano a α.

Si λ es un eigenvalor de A y $\alpha \neq \lambda$, entonces $A - \alpha I$ es invertible si α no es un eigenvalor de A y $1/(\lambda - \alpha)$ es un eigenvalor de $(A - \alpha I)^{-1}$. (Vea el ejercicio 45.) Si α está cerca de λ, entonces $1/(\lambda - \alpha)$ será un eigenvalor dominante de $(A - \alpha I)^{-1}$. De hecho, si α está *muy* cerca de λ, entonces $1/(\lambda - \alpha)$ será *mucho* mayor en magnitud que el siguiente eigenvalor, de modo que (como se anotó en el tercer comentario después del ejemplo 4.31) la convergencia será muy rápida.

Ejemplo 4.34 Use el método de potencia inverso ajustado para aproximar el eigenvalor de

$$A = \begin{bmatrix} 0 & 5 & -6 \\ -4 & 12 & -12 \\ -2 & -2 & 10 \end{bmatrix}$$

que esté más cerca de 5.

Solución Al ajustar, se tiene

$$A - 5I = \begin{bmatrix} -5 & 5 & -6 \\ -4 & 7 & -12 \\ -2 & -2 & 5 \end{bmatrix}$$

Ahora aplique el método de potencia inverso con

$$\mathbf{x}_0 = \mathbf{y}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Resuelva $(A - 5I)\mathbf{x}_1 = \mathbf{y}_0$ para $\mathbf{x}_1$:

$$[A - 5I \mid \mathbf{y}_0] = \left[\begin{array}{ccc|c} -5 & 5 & -6 & 1 \\ -4 & 7 & -12 & 1 \\ -2 & -2 & 5 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 0 & -0.61 \\ 0 & 1 & 0 & -0.88 \\ 0 & 0 & 1 & -0.39 \end{array}\right]$$

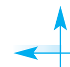
Tabla 4.6

k	0	1	2	3	4	5	6	7
$\mathbf{x}_k$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}-0.61\\-0.88\\-0.39\end{bmatrix}$	$\begin{bmatrix}-0.41\\-0.69\\-0.35\end{bmatrix}$	$\begin{bmatrix}-0.47\\-0.89\\-0.44\end{bmatrix}$	$\begin{bmatrix}-0.49\\-0.95\\-0.48\end{bmatrix}$	$\begin{bmatrix}-0.50\\-0.98\\-0.49\end{bmatrix}$	$\begin{bmatrix}-0.50\\-0.99\\-0.50\end{bmatrix}$	$\begin{bmatrix}-0.50\\-1.00\\-0.50\end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}0.69\\1.00\\0.45\end{bmatrix}$	$\begin{bmatrix}0.59\\1.00\\0.51\end{bmatrix}$	$\begin{bmatrix}0.53\\1.00\\0.50\end{bmatrix}$	$\begin{bmatrix}0.51\\1.00\\0.50\end{bmatrix}$	$\begin{bmatrix}0.50\\1.00\\0.50\end{bmatrix}$	$\begin{bmatrix}0.50\\1.00\\0.50\end{bmatrix}$	$\begin{bmatrix}0.50\\1.00\\0.50\end{bmatrix}$
m_k	1	−0.88	−0.69	−0.89	−0.95	−0.98	−0.99	−1.00

Esto produce

$$\mathbf{x}_1 = \begin{bmatrix}-0.61\\-0.88\\-0.39\end{bmatrix}, \quad m_1 = -0.88 \quad \text{y} \quad \mathbf{y}_1 = \frac{1}{m_1}\mathbf{x}_1 = -\frac{1}{0.88}\begin{bmatrix}-0.61\\-0.88\\-0.39\end{bmatrix} = \begin{bmatrix}0.69\\1\\0.45\end{bmatrix}$$

Continúe de esta forma para obtener los valores en la tabla 4.6, de la cual se deduce que el eigenvalor de A más cercano a 5 es aproximadamente $5 + 1/m_7 \approx 5 + 1/(-1) = 4$, lo que, de hecho, es exacto.



El método de potencia y sus variantes representan solamente un enfoque del cálculo de eigenvalores. En el capítulo 5 se analizará otro método con base en la factorización QR de una matriz. Para un tratamiento más completo de este tema, puede consultar casi cualquier texto acerca de métodos numéricos.

a + bi

Teorema de Gerschgorin

Este teorema se le debe al matemático ruso S. Gerschgorin (1901-1933), quien lo enunció en 1931. No recibió mucha atención sino hasta 1949, cuando lo resucitó Olga Taussky-Todd en una nota que publicó en el *American Mathematical Monthly*.

En esta sección se estudiaron muchas variaciones del método de potencia para aproximar los eigenvalores de una matriz. Todos estos métodos son iterativos y la rapidez con la que convergen depende de la elección del vector inicial. Si sólo tuviera alguna "información confidencial" acerca de la ubicación de los eigenvalores de una matriz dada, entonces podría hacerse una elección juiciosa del vector inicial y acaso acelerar la convergencia del proceso iterativo.

Por fortuna, existe una forma de estimar la ubicación de los eigenvalores de cualquier matriz. El ***teorema del disco de Gerschgorin*** afirma que los eigenvalores de una matriz de $n \times n$ (real o compleja) yacen todos dentro de la unión de n discos circulares en el plano complejo.

Definición Sea $A = [a_{ij}]$ una matriz de $n \times n$ (real o compleja) y sea r_i la suma de los valores absolutos de las entradas fuera de la diagonal en el i-ésimo renglón de A; esto es, $r_i = \sum_{j \neq i} |a_{ij}|$. El ***i-ésimo disco de Gerschgorin*** es el disco circular D_i en el plano complejo con centro a_{ii} y radio r_i. Esto es,

$$D_i = \{z \text{ en } \mathbb{C} : |z - a_{ii}| \leq r_i\}$$

Olga Taussky-Todd (1906-1995) nació en Olmütz en el Imperio Austro-Húngaro (ahora Olmuac, en la República Checa). Recibió su doctorado en teoría de números por parte de la Universidad de Viena en 1930. Durante la Segunda Guerra Mundial trabajó para el Laboratorio Nacional de Física en Londres, donde investigó el problema de la trepidación en la alas de los aviones supersónicos. Aunque el problema involucraba ecuaciones diferenciales, la estabilidad de una aeronave depende de los eigenvalores de una matriz relacionada. Taussky-Todd recordó el teorema de Gerschgorin de sus estudios de posgrado en Viena y pudo usarlos para simplificar los laboriosos cálculos necesarios para determinar de otro modo los eigenvalores relevantes para el problema de la trepidación.

Taussky-Todd se mudó a Estados Unidos en 1947 y diez años después se convirtió en la primera mujer en obtener un nombramiento en California Institute of Technology. En su carrera produjo más de 200 artículos y recibió numerosos premios. Su labor fue importante para el desarrollo de la rama de las matemáticas conocida como teoría de matrices.

Ejemplo 4.35

Bosqueje los discos de Gerschgorin y los eigenvalores para las siguientes matrices:

(a) $A = \begin{bmatrix} 2 & 1 \\ 2 & -3 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & -3 \\ 2 & 3 \end{bmatrix}$

Solución (a) Los dos discos de Gerschgorin tienen centro en 2 y -3 con radios 1 y 2, respectivamente. El polinomio característico de A es $\lambda^2 + \lambda - 8$, de modo que los eigenvalores son

$$\lambda = (-1 \pm \sqrt{1^2 - 4(-8)})/2 \approx 2.37, -3.37$$

La figura 4.15 muestra que los eigenvalores están contenidos dentro de los dos discos de Gerschgorin.

(b) Los dos discos de Gerschgorin tienen centro en 1 y 3, con radios $|-3| = 3$ y 2, respectivamente. El polinomio característico de A es $\lambda^2 - 4\lambda + 9$, de modo que los eigenvalores son

$$\lambda = (4 \pm \sqrt{(-4)^2 - 4(9)})/2 = 2 \pm i\sqrt{5} \approx 2 + 2.23i, 2 - 2.23i$$

La figura 4.16 muestra la ubicación de los eigenvalores en relación con los discos de Gerschgorin.

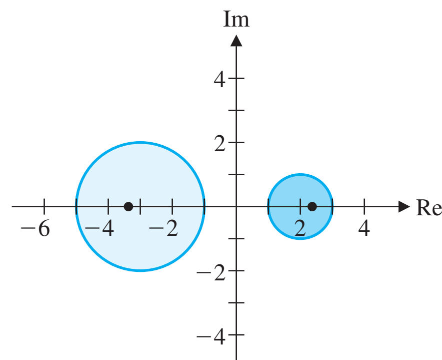


Figura 4.15

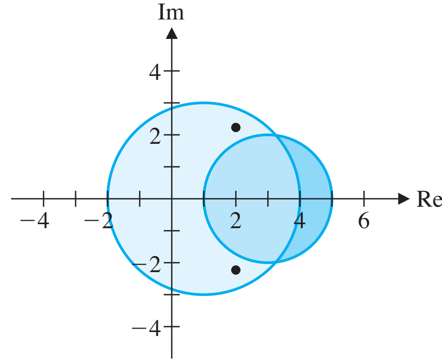
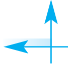


Figura 4.16



Como sugiere el ejemplo 4.35, los eigenvalores de una matriz están contenidos dentro de sus discos de Gershgorin. El siguiente teorema verifica que esto es así.

Teorema 4.29 **Teorema del disco de Gerschgorin**

Sea A una matriz de $n \times n$ (real o compleja). Entonces todo eigenvalor de A está contenido dentro de un disco de Gerschgorin.



Demostración Sea λ un eigenvalor de A con su correspondiente eigenvector $\mathbf{x}$. Sea x_i la entrada de $\mathbf{x}$ con el valor absoluto más grande, y por tanto distinta de cero. (¿Por qué?) Entonces $A\mathbf{x} = \lambda\mathbf{x}$, cuyo i-ésimo renglón es

$$[a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}]\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \lambda x_i \quad \text{o} \quad \sum_{j=1}^{n} a_{ij}x_j = \lambda x_i$$

Al reordenar se tiene

$$(\lambda - a_{ii})x_i = \sum_{j \neq i} a_{ij}x_j \quad \text{o} \quad \lambda - a_{ii} = \frac{\sum_{j \neq i} a_{ij}x_j}{x_i}$$

porque $x_i \neq 0$. Al tomar valores absolutos y usar las propiedades de valor absoluto (vea el Apéndice C), se obtiene

$$|\lambda - a_{ii}| = \left| \frac{\sum_{j \neq i} a_{ij}x_j}{x_i} \right| = \frac{\left| \sum_{j \neq i} a_{ij}x_j \right|}{|x_i|} \leq \frac{\sum_{j \neq i} |a_{ij}x_j|}{|x_i|} = \frac{\sum_{j \neq i} |a_{ij}||x_j|}{|x_i|} \leq \sum_{j \neq i} |a_{ij}| = r_i$$

porque $|x_j| \leq |x_i|$ para $j \neq i$.

Esto establece que el eigenvalor λ está contenido dentro del disco de Gerschgorin con centro en a_{ii} y radio r_i.

Comentarios

- Existe una versión correspondiente del teorema anterior para los discos de Gerschgorin cuyos radios sean la suma de las entradas fuera de la diagonal en la i-ésima *columna* de A.
- Puede demostrarse que si k de los discos de Gerschgorin están separados de los otros discos, entonces exactamente k eigenvalores están contenidos dentro de la unión de dichos k discos. En particular, si un solo disco está separado de los otros, entonces debe contener exactamente un eigenvalor de la matriz. El ejemplo 4.35(a) ilustra esto.
- Note que, en el ejemplo 4.35(a), 0 no está contenido en un disco de Gerschgorin; esto es, 0 no es un eigenvalor de A. Por tanto, sin mayores cálculos, puede deducirse que la matriz A es invertible por el Teorema 4.16. Esta observación es particularmente útil cuando se aplica a matrices más grandes, porque los discos de Gerschgorin pueden determinarse directamente de las entradas de la matriz.

Ejemplo 4.36

Considere la matriz $A = \begin{bmatrix} 2 & 1 & 0 \\ \frac{1}{2} & 6 & \frac{1}{2} \\ 2 & 0 & 8 \end{bmatrix}$. El Teorema de Gerschgorin dice que los eigenvalores de A están contenidos dentro de tres discos con centro en 2, 6 y 8 y radios 1, 1 y 2, respectivamente. Vea la figura 4.17(a). Puesto que el primer disco está separado de los otros dos, debe contener exactamente un eigenvalor, por el segundo Comentario posterior al Teorema 4.29. Dado que el polinomio característico de A tiene coeficientes reales, si tiene raíces complejas (es decir, eigenvalores de A), deben presentarse en pares conjugados. (Vea el Apéndice D.) En consecuencia hay un eigenvalor real único entre 1 y 3, y la unión de los otros dos discos contiene dos eigenvalores (posiblemente complejos) cuyas partes reales se encuentran entre 5 y 10.

Por otra parte, el primer Comentario después del Teorema 4.29 dice que los mismos tres eigenvalores de A están contenidos en discos con centro en 2, 6 y 8 y radios $\frac{5}{2}$, 1 y $\frac{1}{2}$, respectivamente. Vea la figura 4.17(b). Dichos discos están mutuamente separados, de modo que cada uno contiene un solo (y por tanto real) eigenvalor. Al combinar dichos resultados, se deduce que A tiene tres eigenvalores reales, uno en cada uno de los intervalos [1, 3], [5, 7] y [7.5, 8.5]. (Calcule los eigenvalores reales de A para verificar esto.)

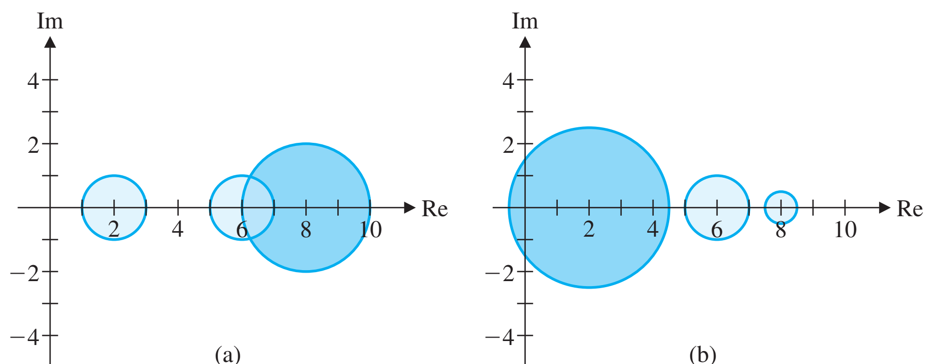


Figura 4.17

CAS

Ejercicios 4.5

En los ejercicios 1-4, se proporciona una matriz A junto con una iteración $\mathbf{x}_5$, *producida como en el ejemplo 4.30.*
(a) Use dichos datos para aproximar un eigenvector dominante cuyo primer componente sea 1 y un correspondiente eigenvalor dominante. (Use precisión a tres lugares decimales.)
(b) Compare su eigenvalor aproximado en el inciso (a) con el eigenvalor dominante real.

1. $A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 4443 \\ 11109 \end{bmatrix}$

2. $A = \begin{bmatrix} 7 & 4 \\ -3 & -1 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 7811 \\ -3904 \end{bmatrix}$

3. $A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 144 \\ 89 \end{bmatrix}$

4. $A = \begin{bmatrix} 1.5 & 0.5 \\ 2.0 & 3.0 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 60.625 \\ 239.500 \end{bmatrix}$

En los ejercicios 5-8, se proporciona una matriz A junto con una iteración $\mathbf{x}_k$, *producida usando el método de potencia, como en el ejemplo 4.31.*
(a) Aproxime el eigenvalor y el eigenvector dominantes al calcular el correspondiente m_k *y* $\mathbf{y}_k$. *(b) Verifique que aproximó un eigenvalor y un eigenvector de A al comparar* $A\mathbf{y}_k$ *con* $m_k\mathbf{y}_k$.

5. $A = \begin{bmatrix} 2 & -3 \\ -3 & 10 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} -3.667 \\ 11.001 \end{bmatrix}$

6. $A = \begin{bmatrix} 5 & 2 \\ 2 & -2 \end{bmatrix}, \mathbf{x}_{10} = \begin{bmatrix} 5.530 \\ 1.470 \end{bmatrix}$

7. $A = \begin{bmatrix} 4 & 0 & 6 \\ -1 & 3 & 1 \\ 6 & 0 & 4 \end{bmatrix}, \mathbf{x}_8 = \begin{bmatrix} 10.000 \\ 0.001 \\ 10.000 \end{bmatrix}$

8. $A = \begin{bmatrix} 1 & 2 & -2 \\ 1 & 1 & -3 \\ 0 & -1 & 1 \end{bmatrix}, \mathbf{x}_{10} = \begin{bmatrix} 3.415 \\ 2.914 \\ -1.207 \end{bmatrix}$

En los ejercicios 9-14, use el método de potencia para aproximar el eigenvalor y el eigenvector dominantes de A. Use el vector inicial dado $\mathbf{x}_0$, *el número específico de iteraciones k y precisión a tres lugares decimales.*

9. $A = \begin{bmatrix} 14 & 12 \\ 5 & 3 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, k = 5$

10. $A = \begin{bmatrix} -6 & 4 \\ 8 & -2 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, k = 6$

11. $A = \begin{bmatrix} 7 & 2 \\ 2 & 3 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, k = 6$

12. $A = \begin{bmatrix} 3.5 & 1.5 \\ 1.5 & -0.5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, k = 6$

13. $A = \begin{bmatrix} 9 & 4 & 8 \\ 4 & 15 & -4 \\ 8 & -4 & 9 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, k = 5$

14. $A = \begin{bmatrix} 3 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 3 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, k = 6$

En los ejercicios 15 y 16, use el método de potencia para aproximar el eigenvalor y el eigenvector dominantes de A con una precisión de dos lugares decimales. Elija cualquier vector inicial que guste (¡pero tenga en mente el primer Comentario de la página 326!) y aplique el método hasta que el dígito en el segundo lugar decimal de las iteraciones deje de cambiar.

15. $A = \begin{bmatrix} 4 & 1 & 3 \\ 0 & 2 & 0 \\ 1 & 1 & 2 \end{bmatrix}$

16. $A = \begin{bmatrix} 12 & 6 & -6 \\ 2 & 0 & -2 \\ -6 & 6 & 12 \end{bmatrix}$

Los cocientes de Rayleigh se describen en el cuarto Comentario de la página 327. En los ejercicios 17-20, vea cómo el método de cocientes de Rayleigh aproxima el eigenvalor dominante más rápidamente que el método de potencia ordinario, calcule los sucesivos cocientes de Rayleigh $R(\mathbf{x}_i)$ *para* $i = 1, \dots, k$ *para la matriz A de los siguientes ejercicios.*

17. Ejercicio 11

18. Ejercicio 12

19. Ejercicio 13

20. Ejercicio 14

Las matrices de los ejercicios 21-24 no son diagonalizables o no tienen un eigenvalor dominante (o ambos). De cualquier forma, aplique el método de potencia con el vector inicial dado $\mathbf{x}_0$ *y realice ocho iteraciones en cada caso. Calcule los eigenvalores y eigenvectores exactos y explique qué ocurre.*

21. $A = \begin{bmatrix} 4 & 1 \\ 0 & 4 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

22. $A = \begin{bmatrix} 3 & 1 \\ -1 & 1 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

23. $A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

24. $A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 25-28, el método de potencia no converge en el eigenvalor y al eigenvector dominantes. Verifique esto usando el vector inicial dado $\mathbf{x}_0$*. Calcule los eigenvalores y eigenvectores exactos y explique qué ocurre.*

25. $A = \begin{bmatrix} -1 & 2 \\ -1 & 1 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

26. $A = \begin{bmatrix} 2 & 1 \\ -2 & 5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

27. $A = \begin{bmatrix} -5 & 1 & 7 \\ 0 & 4 & 0 \\ 7 & 1 & -5 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

28. $A = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}, \mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 29-32, aplique el método de potencia ajustado para aproximar el segundo eigenvalor de la matriz A en el ejercicio dado. Use el vector inicial dado $\mathbf{x}_0$*, k iteraciones y precisión a tres lugares decimales.*

29. Ejercicio 9 **30.** Ejercicio 10

31. Ejercicio 13 **32.** Ejercicio 14

En los ejercicios 33-36, aplique el método de potencia inverso para aproximar, para la matriz A en el ejercicio dado, el eigenvalor que es más pequeño en magnitud. Use el vector inicial dado $\mathbf{x}_0$*, k iteraciones y precisión a tres lugares decimales.*

33. Ejercicio 9 **34.** Ejercicio 10

35. Ejercicio 7, $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, k = 5$

36. Ejercicio 14

En los ejercicios 37-40, use el método de potencia inverso ajustado para aproximar, para la matriz A en el ejercicio dado, el eigenvalor más cercano a α*.*

37. Ejercicio 9, $\alpha = 0$ **38.** Ejercicio 12, $\alpha = 0$

39. Ejercicio 7, $\alpha = 5$ **40.** Ejercicio 13, $\alpha = -2$

El ejercicio 32 de la sección 4.3 demuestra que todo polinomio es (más o menos) el polinomio característico de su propia matriz acompañante. Por tanto, las raíces de un polinomio p son los eigenvalores de C(p). Por tanto, puede usar los métodos de esta sección para aproximar las raíces de cualquier polinomio cuando los resultados exactos no están fácilmente disponibles. En los ejercicios 41-44, aplique el método de potencia inverso ajustado a la matriz acompañante C(p) de p para aproximar la raíz de p más cercana a α *a tres lugares decimales.*

41. $p(x) = x^2 + 2x - 2, \alpha = 0$

42. $p(x) = x^2 - x - 3, \alpha = 2$

43. $p(x) = x^3 - 2x^2 + 1, \alpha = 0$

44. $p(x) = x^3 - 5x^2 + x + 1, \alpha = 5$

45. Sea λ un eigenvalor de A con su correspondiente eigenvector $\mathbf{x}$. Si $\alpha \neq \lambda$ y α no es un eigenvalor de A, demuestre que $1/(\lambda - \alpha)$ es un eigenvalor de $(A - \alpha I)^{-1}$ con su correspondiente eigenvector $\mathbf{x}$. (¿Por qué $A - \alpha I$ debe ser invertible?)

46. Si A tiene un eigenvalor dominante λ_1, demuestre que el eigenespacio E_{λ_1} es unidimensional.

a + bi *En los ejercicios 47-50, dibuje los discos de Gerschgorin para la matriz dada.*

47. $\begin{bmatrix} 1 & 1 & 0 \\ \frac{1}{2} & 4 & \frac{1}{2} \\ 1 & 0 & 5 \end{bmatrix}$ **48.** $\begin{bmatrix} 2 & -i & 0 \\ 1 & 2i & 1+i \\ 0 & 1 & -2i \end{bmatrix}$

49. $\begin{bmatrix} 4-3i & i & 2 & -2 \\ i & -1+i & 0 & 0 \\ 1+i & -i & 5+6i & 2i \\ 1 & -2i & 2i & -5-5i \end{bmatrix}$

50. $\begin{bmatrix} 2 & \frac{1}{2} & 0 & 0 \\ \frac{1}{4} & 4 & \frac{1}{4} & 0 \\ 0 & \frac{1}{6} & 6 & \frac{1}{6} \\ 0 & 0 & \frac{1}{8} & 8 \end{bmatrix}$

51. Una matriz cuadrada es ***estrictamente diagonal dominante*** si el valor absoluto de cada entrada diagonal es mayor que la suma de los valores absolutos de las entradas restantes en dicho renglón. (Vea la sección 2.5.) Use el teorema del disco de Gerschgorin para demostrar que una matriz estrictamente diagonal dominante debe ser invertible. [*Sugerencia:* vea el tercer Comentario después del Teorema 4.29.]

52. Si A es una matriz de $n \times n$, sea $\|A\|$ el máximo de las sumas de los valores absolutos de los renglones de A; esto es $\|A\| = \max_{1 \le i \le n} \left(\sum_{j=1}^{n} |a_{ij}| \right)$. (Vea la sección 7.2.)

Demuestre que, si λ es un eigenvalor de A, entonces $|\lambda| \le \|A\|$.

53. Sea λ un eigenvalor de una matriz estocástica A (vea la sección 3.7). Pruebe que $|\lambda| \le 1$. [*Sugerencia:* aplique el ejercicio 52 a A^T.]

54. Demuestre que los eigenvalores de $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2 & 5 & 0 & 0 \\ \frac{1}{2} & 0 & 3 & \frac{1}{2} \\ 0 & 0 & \frac{3}{4} & 7 \end{bmatrix}$

son todos reales y ubique cada uno de estos eigenvalores dentro de un intervalo cerrado sobre la recta real.

4.6 Aplicaciones y el teorema de Perron-Frobenius

En esta sección se explorarán varias aplicaciones de los eigenvalores y los eigenvectores. Comience por revisar algunas aplicaciones de los capítulos anteriores.

Cadenas de Markov

En la sección 3.7 se introdujeron las cadenas de Markov y se hicieron algunas observaciones acerca de las matrices de transición (estocásticas) asociadas con ellas. En particular, se observó que si P es la matriz de transición de una cadena de Markov, entonces P tiene un vector de estado estacionario $\mathbf{x}$. Esto es, existe un vector $\mathbf{x}$ tal que $P\mathbf{x} = \mathbf{x}$. Esto es equivalente a decir que P tiene 1 como eigenvalor. Ahora está en posición de probar este hecho.

Teorema 4.30 Si P es la matriz de transición de $n \times n$ de una cadena de Markov, entonces 1 es un eigenvalor de P.

Demostración Recuerde que toda matriz de transición es estocástica; por tanto, cada una de sus columnas suma 1. En consecuencia, si $\mathbf{j}$ es un vector renglón que consiste de n números 1, entonces $\mathbf{j}P = \mathbf{j}$. (Vea el ejercicio 13 de la sección 3.7.) Al tomar transpuestas se tiene

$$P^T\mathbf{j}^T = (\mathbf{j}P)^T = \mathbf{j}^T$$

lo que implica que $\mathbf{j}^T$ es un eigenvector de P^T con correspondiente eigenvalor 1. Por el ejercicio 19 de la sección 4.3, P y P^T tienen los mismos eigenvalores, de modo que 1 también es un eigenvalor de P.

De hecho, es verdadero mucho más. Para la mayoría de las matrices de transición, *todo* eigenvalor λ satisface $|\lambda| \leq 1$ y el eigenvalor 1 es *dominante*; esto es, si $\lambda \neq 1$, entonces $|\lambda| < 1$. Son necesarias las siguientes dos definiciones: una matriz se llama ***positiva*** si todas sus entradas son positivas, y una matriz cuadrada se llama ***regular*** si alguna de sus potencias es positiva. Por ejemplo, $A = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$ es positiva, pero $B = \begin{bmatrix} 3 & 1 \\ 2 & 0 \end{bmatrix}$ no lo es. Sin embargo, B es regular, pues $B^2 = \begin{bmatrix} 11 & 3 \\ 6 & 2 \end{bmatrix}$ es positiva.

Teorema 4.31 Sea P una matriz de transición de $n \times n$ con eigenvalor λ.

a. $|\lambda| \leq 1$
b. Si P es regular y $\lambda \neq 1$ entonces $|\lambda| < 1$.

Demostración Como en el Teorema 4.30, el truco para demostrar este teorema es usar el hecho de que P^T tiene los mismos eigenvalores que P.

(a) Sea $\mathbf{x}$ un eigenvector de P^T correspondiente a λ y sea x_k el componente de $\mathbf{x}$ con el valor absoluto más grande m. Entonces $|x_i| \leq |x_k| = m$ para $i = 1, 2, \ldots, n$. Al comparar los k-ésimos componentes de la ecuación $P^T\mathbf{x} = \lambda\mathbf{x}$, se tiene

$$p_{1k}x_1 + p_{2k}x_2 + \cdots + p_{nk}x_n = \lambda x_k$$

(Recuerde que los renglones de P^T son las columnas de P.) Al tomar valores absolutos, se obtiene

$$\begin{aligned} |\lambda| m = |\lambda||x_k| = |\lambda x_k| &= |p_{1k}x_1 + p_{2k}x_2 + \cdots + p_{nk}x_n| \\ &\leq |p_{1k}x_1| + |p_{2k}x_2| + \cdots + |p_{nk}x_n| \\ &= p_{1k}|x_1| + p_{2k}|x_2| + \cdots + p_{nk}|x_n| \\ &\leq p_{1k}m + p_{2k}m + \cdots + p_{nk}m \\ &= (p_{1k} + p_{2k} + \cdots + p_{nk})m = m \end{aligned} \tag{1}$$

La primera desigualdad es consecuencia de la desigualdad del triángulo en $\mathbb{R}$, y la última igualdad proviene del hecho de que los renglones de P^T suman 1. Por tanto, $|\lambda| m \leq m$. Después de dividir entre m, se tiene $|\lambda| \leq 1$, como se deseaba.

(b) Se demostrará la implicación equivalente: si $|\lambda| = 1$, entonces $\lambda = 1$. Primero, se demuestra que es verdadero cuando P (y por tanto P^T) es una matriz positiva. Si $|\lambda| = 1$, entonces todas las desigualdades de las ecuaciones (1) en realidad son igualdades. En particular,

$$p_{1k}|x_1| + p_{2k}|x_2| + \cdots + p_{nk}|x_n| = p_{1k}m + p_{2k}m + \cdots + p_{nk}m$$

De manera equivalente,

$$p_{1k}(m - |x_1|) + p_{2k}(m - |x_2|) + \cdots + p_{nk}(m - |x_n|) = 0 \tag{2}$$

Ahora, dado que P es positiva, $p_{ik} > 0$ para $i = 1, 2, \ldots, n$. Además, $m - |x_i| \geq 0$ para $i = 1, 2, \ldots, n$. Por tanto, cada sumando en la ecuación (2) debe ser cero, y esto sólo puede ocurrir si $|x_i| = m$ para $i = 1, 2, \ldots, n$. Más aún, se tiene la igualdad en la Desigualdad del Triángulo en $\mathbb{R}$ si y sólo si todos los sumandos son positivos o todos son negativos; en otras palabras, que todos los $p_{ik}x_i$ tenga el mismo signo. Esto implica que

$$\mathbf{x} = \begin{bmatrix} m \\ m \\ \vdots \\ m \end{bmatrix} = m\mathbf{j}^T \quad \text{o} \quad \mathbf{x} = \begin{bmatrix} -m \\ -m \\ \vdots \\ -m \end{bmatrix} = -m\mathbf{j}^T$$

donde $\mathbf{j}$ es un vector renglón de n números 1, como en el Teorema 4.30. Por tanto, en cualquier caso, el eigenespacio de P^T correspondiente a λ es $E_\lambda = \text{gen}(\mathbf{j}^T)$.

Pero, al usar la demostración del Teorema 4.30 se ve que $\mathbf{j}^T = P^T\mathbf{j}^T = \lambda\mathbf{j}^T$ y, al comparar componentes, se encuentra que $\lambda = 1$. Esto maneja el caso donde P es positivo.

Si P es regular, entonces alguna potencia de P es positiva, por ejemplo, P^k. Se tiene que P^{k+1} también debe ser positivo. (¿Por qué?) Dado que λ^k y λ^{k+1} son eigenvalores de P^k y P^{k+1}, respectivamente, por el Teorema 4.18 recién se probó que $\lambda^k = \lambda^{k+1} = 1$. Por tanto, $\lambda^k(\lambda - 1) = 0$, lo cual implica que $\lambda = 1$, pues $\lambda = 0$ es imposible si $|\lambda| = 1$.

Ahora es posible explicar parte del comportamiento de las cadenas de Markov que se observaron en el capítulo 3. En el ejemplo 3.64 se vio que, para la matriz de transición

$$P = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$$

y vector de estado inicial $\mathbf{x}_0 = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$, los vectores de estado $\mathbf{x}_k$ convergen en el vector $\mathbf{x} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$, un vector de estado estacionario para P (es decir: $P\mathbf{x} = \mathbf{x}$). A continuación se demostrará que, para cadenas de Markov regulares, esto siempre sucede. De hecho, se de-

mostrará mucho más. Recuerde que los vectores de estado $\mathbf{x}_k$ satisfacen $\mathbf{x}_k = P^k\mathbf{x}_0$. Investigue lo que ocurre a las potencias P^k cuando K se vuelve grande.

Ejemplo 4.37

La matriz de transición $P = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$ tiene ecuación característica

$$0 = \det(P - \lambda I) = \begin{vmatrix} 0.7 - \lambda & 0.2 \\ 0.3 & 0.8 - \lambda \end{vmatrix} = \lambda^2 - 1.5\lambda + 0.5 = (\lambda - 1)(\lambda - 0.5)$$

de modo que sus eigenvalores son $\lambda_1 = 1$ y $\lambda_2 = 0.5$. (Note que, gracias a los Teoremas 4.30 y 4.31, se sabía por anticipado que 1 sería un eigenvalor y el otro eigenvalor sería menor que 1 en valor absoluto. Sin embargo, todavía es necesario calcular λ_2.) Los eigenespacios son

$$E_1 = \text{gen}\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) \quad \text{y} \quad E_{0.5} = \text{gen}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$$

De este modo, al tomar $Q = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}$, se sabe que $Q^{-1}PQ = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix} = D$. De acuerdo con el método utilizado en el ejemplo 4.29 de la sección 4.4, se tiene

$$P^k = QD^kQ^{-1} = \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 1^k & 0 \\ 0 & (0.5)^k \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}^{-1}$$

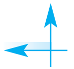
Ahora, a medida que $k \to \infty$, $(0.5)^k \to 0$, de modo que

$$D^k \to \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad P^k \to \begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 3 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 0.4 & 0.4 \\ 0.6 & 0.6 \end{bmatrix}$$

(Observe que las columnas de esta "matriz límite" son idénticas y cada una es un vector de estado estacionario para P.) Ahora sea $\mathbf{x}_0 = \begin{bmatrix} a \\ b \end{bmatrix}$ cualquier vector de probabilidad inicial (es decir: $a + b = 1$). Entonces

$$\mathbf{x}_k = P^k\mathbf{x}_0 \to \begin{bmatrix} 0.4 & 0.4 \\ 0.6 & 0.6 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.4a + 0.4b \\ 0.6a + 0.6b \end{bmatrix} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$$

No sólo explica lo que se vio en el ejemplo 3.64, también dice que los vectores de estado $\mathbf{x}_k$ ¡convergen en el vector de estado estacionario $\mathbf{x} = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$ para *cualquier* elección de $\mathbf{x}_0$!



No hay nada especial en el ejemplo 4.37. El siguiente teorema demuestra que este tipo de comportamiento *siempre* ocurre con matrices de transición regulares. Antes de poder presentar el teorema, es necesario el siguiente lema.

Lema 4.32

Sea P una matriz de transición regular de $n \times n$. Si P es diagonalizable, entonces el eigenvalor dominante $\lambda_1 = 1$ tiene multiplicidad algebraica 1.

Demostración Los eigenvalores de P y P^T son iguales. A partir de la demostración del Teorema 4.31(b), $\lambda_1 = 1$ tiene multiplicidad geométrica 1 como un eigenvalor de P^T. Puesto que P es diagonalizable, también lo es P^T, por el ejercicio 41 de la sección 4.4. En consecuencia, el eigenvalor $\lambda_1 = 1$ tiene multiplicidad algebraica 1, por el teorema de diagonalización.

Teorema 4.33 Sea P una matriz de transición de $n \times n$ regular. Entonces, cuando $k \to \infty$, P^k tiende a una matriz L de $n \times n$ cuyas columnas son idénticas, cada una igual al mismo vector $\mathbf{x}$. Este vector $\mathbf{x}$ es un vector de probabilidad de estado estacionario para P.

Vea *Finite Markov Chains* de J. G. Kemeny y J. L. Snell (Nueva York: Springer-Verlag, 1976).

Demostración Para simplificar la demostración, sólo se considerará el caso donde P es diagonalizable. Sin embargo, el teorema es verdadero sin esta suposición.

Diagonalice P como $Q^{-1}PQ = D$ o, de manera equivalente, $P = QDQ^{-1}$, donde

$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

A partir de los Teoremas 4.30 y 4.31, se sabe que cada eigenvalor λ_i es 1 o satisface $|\lambda_i| < 1$. Por tanto, conforme $k \to \infty$, λ_i^k tiende a 1 o 0 para $i = 1, \ldots, n$. Se tiene que D^k tiende a una matriz diagonal, por decir, D^*, donde cada una de sus entradas diagonales es 1 o 0. Por tanto, $P^k = QD^kQ^{-1}$ tiende a $L = QD^*Q^{-1}$. Escriba

$$\lim_{k\to\infty} P^k = L$$

Se tomaron algunas libertades con la noción de límite. No obstante, dichos pasos deben ser intuitivamente claros. Las demostraciones rigurosas a partir de las propiedades de los límites pueden encontrarse en un curso de cálculo. En lugar de desviarse con una discusión de los límites de las matrices, se omitirán las demostraciones.

Observe que

$$PL = P\lim_{k\to\infty} P^k = \lim_{k\to\infty} PP^k = \lim_{k\to\infty} P^{k+1} = L$$

Por tanto, cada columna de L es un eigenvector de P correspondiente a $\lambda_1 = 1$. Para ver que cada una de dichas columnas es un vector de *probabilidad* (es decir, L es una matriz estocástica), sólo es necesario observar que, si $\mathbf{j}$ es el vector renglón con n números 1, entonces

$$\mathbf{j}L = \mathbf{j}\lim_{k\to\infty} P^k = \lim_{k\to\infty} \mathbf{j}P^k = \lim_{k\to\infty} \mathbf{j} = \mathbf{j}$$

pues P^k es una matriz estocástica, por el ejercicio 14 en la sección 3.7. El ejercicio 13 de la sección 3.7 implica ahora que L es estocástica.

Sólo es necesario mostrar que las columnas de L son idénticas. La i-ésima columna de L es $L\mathbf{e}_i$, donde $\mathbf{e}_i$ es el i-ésimo vector base estándar. Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ eigenvectores de P que forman una base de $\mathbb{R}^n$, con $\mathbf{v}_1$ correspondiente a $\lambda_1 = 1$. Escriba

$$\mathbf{e}_i = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$$

para escalares $c_1, c_2, \ldots, c_n$. Entonces, por el comentario en el recuadro después del ejemplo 4.21,

$$P^k\mathbf{e}_i = c_1 1^k\mathbf{v}_1 + c_2\lambda_2^k\mathbf{v}_2 + \cdots + c_n\lambda_n^k\mathbf{v}_n$$

Por el Lema 4.32, $\lambda_j \neq 1$ para $j \neq 1$, de modo que, por el Teorema 4.31(b), $|\lambda_j| < 1$ para $j \neq 1$. En consecuencia, $\lambda_j^k \to 0$ conforme $k \to \infty$, para $j \neq 1$. Se tiene que

$$L\mathbf{e}_i = \lim_{k\to\infty} P^k\mathbf{e}_i = c_1\mathbf{v}_1$$

En otras palabras, la columna i de L es un eigenvector que corresponde a $\lambda_1 = 1$. Pero se demostró que las columnas de L probablemente son vectores, de modo que $L\mathbf{e}_i$ es el *único* múltiplo $\mathbf{x}$ de $\mathbf{v}_1$ cuyos componentes suman 1. Dado que esto es cierto para cada columna de L, implica que todas las columnas de L son idénticas, cada una igual a este vector $\mathbf{x}$.

Comentario Dado que L es una matriz estocástica, puede interpretarla como la ***matriz de transición de rango largo*** de la cadena de Markov. Esto es, L_{ij} representa la probabilidad de estar en el estado i, habiendo partido del estado j, *si las transiciones continuaran de manera infinita*. El hecho de que las columnas de L sean idénticas dice que *no importa el estado de partida*, como lo ilustra el siguiente ejemplo.

Ejemplo 4.38

Recuerde la rata en la caja del ejemplo 3.65. La matriz de transición era

$$P = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 0 \end{bmatrix}$$

Se determinó que el vector de probabilidad de estado estacionario es

$$\mathbf{x} = \begin{bmatrix} \frac{1}{4} \\ \frac{3}{8} \\ \frac{3}{8} \end{bmatrix}$$

Por tanto, las potencias de P tienden a

$$L = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{3}{8} & \frac{3}{8} & \frac{3}{8} \\ \frac{3}{8} & \frac{3}{8} & \frac{3}{8} \end{bmatrix} = \begin{bmatrix} 0.250 & 0.250 & 0.250 \\ 0.375 & 0.375 & 0.375 \\ 0.375 & 0.375 & 0.375 \end{bmatrix}$$

de donde se puede ver que la rata *a la larga* pasará 25% de su tiempo en el compartimiento 1 y 37.5% de su tiempo en cada uno de los otros dos compartimientos.

La discusión acerca de las cadenas de Markov concluye al demostrar que el vector de estado estacionario $\mathbf{x}$ es independiente del estado inicial. La demostración se adapta con facilidad para incluir el caso de los vectores de estado cuyos componentes suman una constante arbitraria, por ejemplo, s. En los ejercicios, se le pide demostrar algunas otras propiedades de las cadenas de Markov regulares.

Teorema 4.34

Sea P una matriz de transición de $n \times n$ regular, con $\mathbf{x}$ como el vector de probabilidad de estado estacionario para P, como en el Teorema 4.33. Entonces, para cualquier vector de probabilidad inicial $\mathbf{x}_0$, la sucesión de iteraciones de $\mathbf{x}_k$ tiende a $\mathbf{x}$.

Demostración Sea

$$\mathbf{x}_0 = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

donde $x_1 + x_2 + \cdots + x_n = 1$. Dado que $\mathbf{x}_k = P^k\mathbf{x}_0$, debe demostrar que $\lim_{k\to\infty} P^{k\,k}\mathbf{x}_0 = \mathbf{x}$. Ahora, por el Teorema 4.33, la matriz de transición de rango largo es $L = [\mathbf{x}\ \ \mathbf{x}\ \ \cdots\ \ \mathbf{x}]$ y $\lim_{k\to\infty} P^k = L$. Por tanto

$$\begin{aligned}\lim_{k\to\infty} P^k\mathbf{x}_0 &= (\lim_{k\to\infty} P^k)\mathbf{x}_0 = L\mathbf{x}_0\\ &= [\mathbf{x}\quad \mathbf{x}\quad \cdots\quad \mathbf{x}]\begin{bmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix}\\ &= x_1\mathbf{x} + x_2\mathbf{x} + \cdots + x_n\mathbf{x}\\ &= (x_1 + x_2 + \cdots + x_n)\mathbf{x} = \mathbf{x}\end{aligned}$$

Crecimiento poblacional

Regrese al modelo de Leslie de crecimiento poblacional, que se exploró por primera vez en la sección 3.7. En el ejemplo 3.67 de dicha sección se vio que, para la matriz de Leslie

$$L = \begin{bmatrix} 0 & 4 & 3\\ 0.5 & 0 & 0\\ 0 & 0.25 & 0\end{bmatrix}$$

las iteraciones de los vectores de población comienzan a aproximarse a un múltiplo del vector

$$\mathbf{x} = \begin{bmatrix} 18\\ 6\\ 1\end{bmatrix}$$

En otras palabras, las tres clases etáreas de esta población finalmente terminarán en la razón 18 : 6 : 1. Más aún, una vez alcanzado este estado, es estacionario, pues las razones para los años siguientes están dadas por

$$L\mathbf{x} = \begin{bmatrix} 0 & 4 & 3\\ 0.5 & 0 & 0\\ 0 & 0.25 & 0\end{bmatrix}\begin{bmatrix} 18\\ 6\\ 1\end{bmatrix} = \begin{bmatrix} 27\\ 9\\ 1.5\end{bmatrix} = 1.5\mathbf{x}$$

y los componentes todavía están en la proporción $27{:}9{:}1.5 = 18{:}6{:}1$. Observe que 1.5 representa la *tasa de crecimiento* de esta población cuando alcanza su estado estacionario.

Ahora puede reconocer que $\mathbf{x}$ es un eigenvector de L que corresponde al eigenvalor $\lambda = 1.5$. Por tanto, la tasa de crecimiento de estado estacionario es un eigenvalor *positivo* de L, y un eigenvector que corresponde a este eigenvalor representa los tamaños *relativos* de las clases etáreas cuando se alcanza el estado estacionario. Puede calcular esto directamente sin tener que iterar como se hizo antes.

Ejemplo 4.39 Encuentre la tasa de crecimiento de estado estacionario y las correspondientes razones entre las clases etáreas para la anterior matriz de Leslie L.

Solución Necesita encontrar todos los eigenvalores positivos y sus correspondientes eigenvectores de L. El polinomio característico de L es

$$\det(L - \lambda I) = \begin{vmatrix} -\lambda & 4 & 3\\ 0.5 & -\lambda & 0\\ 0 & 0.25 & -\lambda\end{vmatrix} = -\lambda^3 + 2\lambda + 0.375$$

así que debe resolver $-\lambda^3 + 2\lambda + 0.375 = 0$ o, de manera equivalente, $8\lambda^3 - 16\lambda - 3 = 0$. Al factorizar, se tiene

$$(2\lambda - 3)(4\lambda^2 + 6\lambda + 1) = 0$$

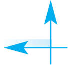
(Vea el Apéndice D.) Puesto que el segundo factor sólo tiene las raíces $(-3 + \sqrt{5})/4 \approx -0.19$ y $(-3 - \sqrt{5})/4 \approx -1.31$, la única raíz positiva de esta ecuación es $\lambda = \frac{3}{2} = 1.5$ Los correspondientes eigenvectores están en el espacio nulo de $L - 1.5I$, que se encontró mediante reducción por renglón:

$$[L - 1.5I \mid \mathbf{0}] = \left[\begin{array}{ccc|c} -1.5 & 4 & 3 & 0 \\ 0.5 & -1.5 & 0 & 0 \\ 0 & 0.25 & -1.5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -18 & 0 \\ 0 & 1 & -6 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Por tanto, si $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ es un eigenvector correspondiente a $\lambda = 1.5$, satisface $x_1 = 18x_3$ y $x_2 = 6x_3$. Esto es,

$$E_{1.5} = \left\{ \begin{bmatrix} 18x_3 \\ 6x_3 \\ x_3 \end{bmatrix} \right\} = \text{gen}\left(\begin{bmatrix} 18 \\ 6 \\ 1 \end{bmatrix} \right)$$

En consecuencia, la tasa de crecimiento de estado estacionario es 1.5 y, cuando se alcanza esta tasa, las clases etáreas están en la razón 18 : 6 : 1, como se vio antes.



En el ejemplo 4.39 sólo hubo un candidato para la tasa de crecimiento de estado estacionario: el eigenvalor positivo único de L. Pero, ¿qué hubiera hecho si L tuviera más de un eigenvalor positivo o ninguno? Aparentemente es muy afortunado que haya un eigenvector correspondiente cuyos componentes son todos positivos, lo que permite relacionar dichos componentes con el tamaño de la población. Puede demostrar que esta situación no es accidental; esto es, *toda* matriz de Leslie tiene exactamente un eigenvalor positivo y un correspondiente eigenvector con componentes positivos.

Recuerde que la forma de una matriz de Leslie es

$$L = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_{n-1} & b_n \\ s_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & s_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & s_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & s_{n-1} & 0 \end{bmatrix} \qquad (3)$$

Dado que las entradas s_j representan probabilidades de supervivencia, se supondrá que todas son distintas de cero (de otro modo, la población moriría rápidamente). También se supondrá que al menos uno de los parámetros de natalidad b_i es distinto de cero (de otro modo no habría nacimientos y, nuevamente, la población moriría). Con estas suposiciones planteadas, ahora puede probar como teorema la aseveración hecha anteriormente.

Teorema 4.35 Toda matriz de Leslie tiene un eigenvalor positivo único y un eigenvector correspondiente con componentes positivos.

Demostración Sea L como en la ecuación (3). El polinomio característico de L es

$$\begin{aligned} c_L(\lambda) &= \det(L - \lambda I) \\ &= (-1)^n(\lambda^n - b_1\lambda^{n-1} - b_2s_1\lambda^{n-2} - b_3s_1s_2\lambda^{n-3} - \cdots - b_ns_1s_2\cdots s_{n-1}) \\ &= (-1)^n f(\lambda) \end{aligned}$$

(En el ejercicio 16 se le pide probar esto.) Por tanto, los eigenvalores de L son las raíces de $f(\lambda)$. Dado que al menos uno de los parámetros de natalidad b_i es positivo y todas las probabilidades de supervivencia s_j son positivas, los coeficientes de $f(\lambda)$ cambian de signo exactamente una vez. En consecuencia, por la regla de los signos de Descartes (Apéndice D), $f(\lambda)$ tiene exactamente una raíz positiva. Llámela λ_1.

Mediante cálculo directo, puede comprobar que un eigenvector que corresponde a λ_1 es

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ s_1/\lambda_1 \\ s_1s_2/\lambda_1^2 \\ s_1s_2s_3/\lambda_1^3 \\ \vdots \\ s_1s_2s_3\cdots s_{n-1}/\lambda_1^{n-1} \end{bmatrix}$$

(En el ejercicio 18 se le pide demostrar esto.) Claramente, todos los componentes de $\mathbf{x}_1$ son positivos.

De hecho, hay algo más de verdad. Con el requisito adicional de que *dos* parámetros de natalidad *consecutivos* b_i y b_{i+1} sean positivos, se evidencia que el único eigenvalor positivo λ_1 de L es *dominante*; esto es, cualquier otro eigenvalor λ (real o complejo) de L satisface $|\lambda| < \lambda_1$. (Está más allá del ámbito de este libro demostrar este resultado, pero en el ejercicio 27 se bosqueja una demostración parcial para los lectores familiarizados con el álgebra de números complejos.) Esto explica por qué se consigue convergencia en un vector de estado estacionario cuando se iteran los vectores de población: ¡es justo el método de potencia que trabaja para usted!

El Teorema de Perron-Frobenius

En las dos aplicaciones anteriores, cadenas de Markov y matrices de Leslie, se vio que el eigenvalor de interés era positivo y dominante. Más aún, hubo un correspondiente eigenvector con componentes positivos. Se evidencia que un teorema notable garantiza que este será el caso para una gran clase de matrices, incluidas muchas de las que se han considerado. La primera versión de este teorema es para matrices positivas.

Primero, es necesaria cierta terminología y notación. Recuerde referirse a un vector como *positivo* si todos sus componentes son positivos. Para dos matrices de $A = [a_{ij}]$ $m \times n$ y $B = [b_{ij}]$, se escribirá $A \geq B$ si $a_{ij} \geq b_{ij}$ para todo i y j. (Definiciones similares se aplicarán para $A > B$, $A \leq B$, etcétera.) Por tanto, un vector positivo $\mathbf{x}$ satisface $\mathbf{x} > \mathbf{0}$. Defina $|A| = [|a_{ij}|]$ como la matriz de los valores absolutos de las entradas de A.

Oskar Perron (1880-1975) fue un matemático alemán que trabajó en muchos campos de las matemáticas, incluidos análisis, ecuaciones diferenciales, álgebra, geometría y teoría de números. El teorema de Perron se publicó en 1907 en un ensayo acerca de fracciones continuas.

Teorema 4.36 **Teorema de Perron**

Sea A una matriz positiva de $n \times n$. Entonces A tiene un eigenvalor real λ_1 con las siguientes propiedades:

a. $\lambda_1 > 0$
b. λ_1 tiene un eigenvector positivo correspondiente.
c. Si λ es cualquier otro eigenvalor de A, entonces $|\lambda| \leq \lambda_1$.

Intuitivamente, puede ver por qué deben ser verdaderos los dos primeros enunciados. Considere el caso de una matriz positiva A de 2×2. La correspondiente transformación matricial mapea el primer cuadrante del plano adecuadamente en sí mismo, pues todos los componentes son positivos. Si repetidamente se permite a A actuar sobre las imágenes que se tienen, necesariamente convergen hacia algún rayo en el primer cuadrante (figura 4.18). Un vector director para este rayo será un vector positivo $\mathbf{x}$, que debe mapearse en algún múltiplo positivo de sí mismo (por decir, λ_1), pues A deja el rayo fijo. En otras palabras, $A\mathbf{x} = \lambda_1\mathbf{x}$, con $\mathbf{x}$ y λ_1 ambos positivos.

Demostración Para algunos vectores distintos de cero $\mathbf{x}$, $A\mathbf{x} \geq \lambda\mathbf{x}$ para algún escalar λ. Cuando esto sucede, entonces $A(k\mathbf{x}) \geq \lambda(k\mathbf{x})$ para toda $k > 0$; por tanto, sólo es necesario considerar vectores *unitarios* $\mathbf{x}$. En el capítulo 7 se verá que A mapea al conjunto de todos los vectores unitarios en $\mathbb{R}^n$ (la *esfera unitaria*) en un "elipsoide generalizado". De este modo, conforme $\mathbf{x}$ varía sobre los vectores no negativos en esta esfera unitaria, habrá un valor máximo de λ tal que $A\mathbf{x} \geq \lambda\mathbf{x}$. (Vea la figura 4.19.) Denote este número mediante λ_1 y el correspondiente vector unitario mediante $\mathbf{x}_1$.

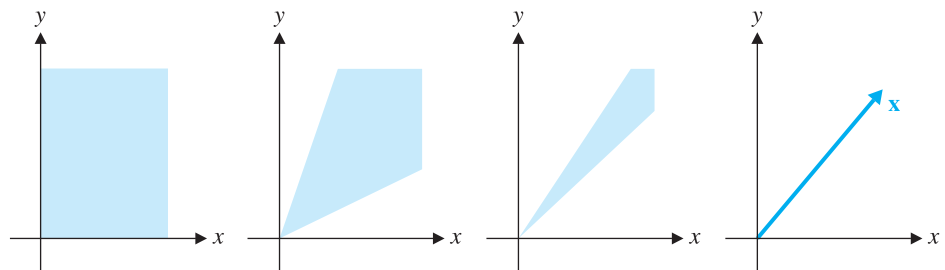


Figura 4.18

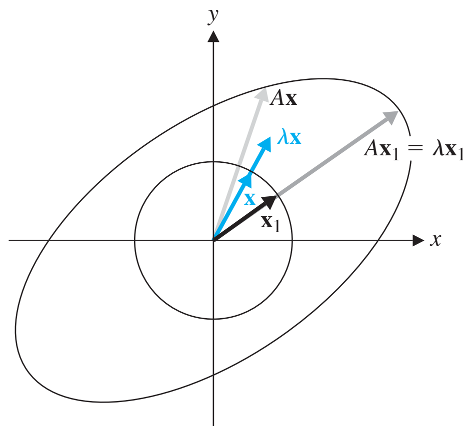


Figura 4.19

Ahora se demostrará que $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$. Si no, entonces $A\mathbf{x}_1 > \lambda_1\mathbf{x}_1$, y al aplicar A nuevamente, se obtiene

$$A(A\mathbf{x}_1) > A(\lambda_1\mathbf{x}_1) = \lambda_1(A\mathbf{x}_1)$$

donde la desigualdad se conserva, pues A es positiva. (Vea el ejercicio 40 y la sección 3.7, ejercicio 36.) Pero entonces $\mathbf{y} = (1/\|A\mathbf{x}_1\|)A\mathbf{x}_1$ es un vector unitario que satisface $A\mathbf{y} > \lambda_1\mathbf{y}$, de modo que habrá algún $\lambda_2 > \lambda_1$ tal que $A\mathbf{y} \geq \lambda_2\mathbf{y}$. Esto contradice el hecho de que λ_1 era el valor máximo con esta propiedad. En consecuencia, debe ser el caso que $A\mathbf{x}_1 = \lambda_1\mathbf{x}_1$; esto es: λ_1 es un eigenvalor de A.

Ahora A es positivo y $\mathbf{x}_1$ es positivo, de modo que $\lambda_1\mathbf{x}_1 = A\mathbf{x}_1 > \mathbf{0}$. Esto significa que $\lambda_1 > 0$ y $\mathbf{x}_1 > \mathbf{0}$, lo que completa la demostración de (a) y (b).

Para demostrar (c), suponga que λ es cualquier otro eigenvalor (real o complejo) de A con correspondiente eigenvector $\mathbf{z}$. Entonces $A\mathbf{z} = \lambda\mathbf{z}$ y, al tomar valores absolutos, se tiene

$$A|\mathbf{z}| = |A||\mathbf{z}| \geq |A\mathbf{z}| = |\lambda\mathbf{z}| = |\lambda||\mathbf{z}| \tag{4}$$

donde la desigualdad intermedia surge de la Desigualdad del Triángulo. (Vea el ejercicio 40.) Dado que $|\mathbf{z}| > \mathbf{0}$, el vector unitario $\mathbf{u}$ en la dirección de $|\mathbf{z}|$ también es positivo y satisface $A\mathbf{u} \geq |\lambda|\mathbf{u}$. Por ser máxima de λ_1 desde la primera parte de esta demostración, se debe tener $|\lambda| \leq \lambda_1$.

De hecho, hay algo más de verdad. Se evidencia que λ_1 es dominante, de modo que $|\lambda| < \lambda_1$ para cualquier eigenvalor $\lambda \neq \lambda_1$. También es el caso que λ_1 tiene multiplicidad algebraica y por tanto geométrica de 1. No se demostrarán estos hechos.

El teorema de Perron puede generalizarse a partir de matrices positivas a ciertas matrices no negativas. Frobenius lo hizo en 1912. El resultado requiere una condición técnica en la matriz. Una matriz cuadrada A se llama ***reducible*** si, sujeta a alguna permutación de los renglones y la misma permutación de las columnas, A puede escribirse en forma de bloque como

$$\begin{bmatrix} B & C \\ O & D \end{bmatrix}$$

donde B y D son cuadradas. De manera equivalente, A es reducible si hay alguna matriz permutación P tal que

$$PAP^T = \begin{bmatrix} B & C \\ O & D \end{bmatrix}$$

(Vea la página 193.) Por ejemplo, la matriz

$$A = \begin{bmatrix} 2 & 0 & 0 & 1 & 3 \\ 4 & 2 & 1 & 5 & 5 \\ 1 & 2 & 7 & 3 & 0 \\ 6 & 0 & 0 & 2 & 1 \\ 1 & 0 & 0 & 7 & 2 \end{bmatrix}$$

es reducible, pues el intercambio de los renglones 1 y 3 y luego las columnas 1 y 3 produce

$$\left[\begin{array}{cc|ccc} 7 & 2 & 1 & 3 & 0 \\ 1 & 2 & 4 & 5 & 5 \\ \hline 0 & 0 & 2 & 1 & 3 \\ 0 & 0 & 6 & 2 & 1 \\ 0 & 0 & 1 & 7 & 2 \end{array}\right]$$

(Esto es justo PAP^T, donde

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$


¡Compruebe esto!)

Una matriz cuadrada A que no es reducible se llama ***irreductible***. Si $A^k > O$ para alguna k, entonces A se llama ***primitiva***. Por ejemplo, toda cadena de Markov regular tiene una matriz de transición primitiva, por definición. No es difícil demostrar que toda matriz primitiva es irreductible. (¿Puede ver por qué? Trate de demostrar la contrapuesta de esto.)


Teorema 4.37 **El teorema de Perron-Frobenius**

Sea A una matriz de $n \times n$ no negativa irreductible. Entonces A tiene un eigenvalor real λ_1 con las siguientes propiedades:

a. $\lambda_1 > 0$
b. λ_1 tiene un correspondiente eigenvector positivo.
c. Si λ es cualquier otro eigenvalor de A, entonces $|\lambda| \leq \lambda_1$. Si A es primitiva, entonces esta desigualdad es estricta.
d. Si λ es un eigenvalor de A tal que $|\lambda| = \lambda_1$, entonces λ es una raíz (compleja) de la ecuación $\lambda^n - \lambda_1^n = 0$.
e. λ_1 tiene multiplicidad algebraica de 1.

Vea *Matrix Analysis* de R. A. Horn y C. R. Johnson (Cambridge, Inglaterra: Cambridge University Press, 1985).

El lector interesado puede encontrar una prueba del teorema de Perron-Frobenius en muchos textos acerca de matrices no negativas o análisis de matrices. El eigenvalor λ_1 con frecuencia se conoce como ***raíz de Perron*** de A, y un correspondiente vector *de probabilidad* (que necesariamente es único) se llama ***eigenvector de Perron*** de A.

Relaciones de recurrencia lineal

Los ***números de Fibonacci*** son los que forman la sucesión 0, 1, 1, 2, 3, 5, 8, 13, 21, . . . , donde, después de los primeros dos términos, cada nuevo término se obtiene al sumar los dos términos que lo anteceden. Si el n-ésimo número de Fibonacci se denota f_n, entonces esta sucesión se define completamente mediante las ecuaciones $f_0 = 0$, $f_1 = 1$ y, para $n \geq 2$,

$$f_n = f_{n-1} + f_{n-2}$$

Esta última ecuación es un ejemplo de una relación de recurrencia lineal. Se regresará a los números de Fibonacci, pero primero se considerarán las relaciones de recurrencia lineal de manera un poco más general.

Leonardo de Pisa (1170-1250), es mejor conocido por su sobrenombre, Fibonacci, que significa "hijo de Bonaccio". Escribió algunos libros importantes, muchos de los cuales han sobrevivido, incluido *Liber abaci* y *Liber quadratorum*. La sucesión de Fibonacci aparece como la solución a un problema en *Liber abaci*: "Cierto hombre puso un par de conejos en un lugar rodeado en todos lados por una pared. ¿Cuántos pares de conejos pueden producirse en un año a partir de este par, si supone que cada mes cada pareja engendra un nuevo par que a partir del segundo mes se vuelve productivo?" El matemático francés Edouard Lucas (1842-1891) dio el nombre *números de Fibonacci* a los términos de esta sucesión.

Definición Sea $(x_n) = (x_0, x_1, x_2, \ldots)$ una sucesión de números que se define del modo siguiente:

1. $x_0 = a_0, x_1 = a_1, \ldots, x_{k-1} = a_{k-1}$, donde $a_0, a_1, \ldots, a_{k-1}$ son escalares.
2. Para todo $n \geq k$, $x_n = c_1x_{n-1} + c_2x_{n-2} + \cdots + c_kx_{n-k}$, donde $c_1, c_2, \ldots, c_k$ son escalares.

Si $c_k \neq 0$, la ecuación en (2) se llama ***relación de recurrencia lineal de orden k***. Las ecuaciones en (1) se conocen como ***condiciones iniciales*** de la recurrencia.

Por tanto, los números de Fibonacci satisfacen una relación de recurrencia lineal de orden 2.

Comentarios

- Si, con la finalidad de definir el n-ésimo término en una relación de recurrencia, se requiere el $(n - k)$-ésimo término pero ningún término antes que él, entonces la relación tiene orden k.
- El número de condiciones iniciales es el orden de la relación de recurrencia.
- No es necesario que el primer término de la sucesión se llame x_0. Podría comenzar en x_1 o en cualquier otro.
- Es posible tener incluso relaciones de recurrencia lineal más generales al permitir que los coeficientes c_i sean *funciones* en lugar de escalares al permitir un coeficiente adicional aislado, que también puede ser una función. Un ejemplo sería la recurrencia

$$x_n = 2x_{n-1} - n^2x_{n-2} + \frac{1}{n}x_{n-3} + n$$

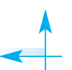
Aquí no se considerarán tales recurrencias.

Ejemplo 4.40 Considere la sucesión (x_n) definida por las condiciones iniciales $x_1 = 1, x_2 = 5$ y la relación de recurrencia $x_n = 5x_{n-1} - 6x_{n-2}$ para $n \geq 2$. Escriba los primeros cinco términos de esta sucesión.

Solución Se proporcionan los dos primeros términos. La relación de recurrencia se usa para calcular los siguientes tres términos. Se tiene

$$x_3 = 5x_2 - 6x_1 = 5 \cdot 5 - 6 \cdot 1 = 19$$
$$x_4 = 5x_3 - 6x_2 = 5 \cdot 19 - 6 \cdot 5 = 65$$
$$x_5 = 5x_4 - 6x_3 = 5 \cdot 65 - 6 \cdot 19 = 211$$

de modo que la sucesión comienza 1, 5, 19, 65, 211, ...



Claramente, si se está interesado en, por ejemplo, el 100-ésimo término de la sucesión en el ejemplo 4.40, entonces el enfoque usado sería más bien tedioso, pues tendría que aplicar la relación de recurrencia 98 veces. Sería agradable si pudiera encontrar una fórmula *explícita* para x_n en función de n. Al hecho de encontrar tal término se le llamará ***resolver*** la relación de recurrencia. El proceso se ilustrará con la sucesión del ejemplo 4.40.

Para comenzar, reescriba la relación de recurrencia como una ecuación matricial. Sea

$$A = \begin{bmatrix} 5 & -6 \\ 1 & 0 \end{bmatrix}$$

e introduzca los vectores $\mathbf{x}_n = \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix}$ para $n \geq 2$. Por tanto, $\mathbf{x}_2 = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} x_3 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 5 \end{bmatrix}$, $\mathbf{x}_4 = \begin{bmatrix} x_4 \\ x_3 \end{bmatrix} = \begin{bmatrix} 65 \\ 19 \end{bmatrix}$, etcétera. Ahora observe que, para $n \geq 2$, se tiene

$$A\mathbf{x}_{n-1} = \begin{bmatrix} 5 & -6 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_{n-1} \\ x_{n-2} \end{bmatrix} = \begin{bmatrix} 5x_{n-1} - 6x_{n-2} \\ x_{n-1} \end{bmatrix} = \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} = \mathbf{x}_n$$

Note que este es el mismo tipo de ecuación que se encontró con las cadenas de Markov y las matrices de Leslie. Como en dichos casos, puede escribir

$$\mathbf{x}_n = A\mathbf{x}_{n-1} = A^2\mathbf{x}_{n-2} = \cdots = A^{n-2}\mathbf{x}_2$$

Ahora use la técnica del ejemplo 4.29 para calcular las potencias de A.

La ecuación característica de A es

$$\lambda^2 - 5\lambda + 6 = 0$$

de la cual se encuentra que los eigenvalores son $\lambda_1 = 3$ y $\lambda_2 = 2$. (Note que la forma de la ecuación característica sigue la de la relación de recurrencia. Si escribe la recurrencia como $x_n - 5x_{n-1} + 6x_{n-2} = 0$, ¡es evidente que los coeficientes son exactamente iguales!) Los correspondientes eigenespacios son

$$E_3 = \text{gen}\left(\begin{bmatrix} 3 \\ 1 \end{bmatrix}\right) \quad \text{y} \quad E_2 = \text{gen}\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right)$$

Al hacer $P = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}$, se sabe que $P^{-1}AP = D = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$. Entonces $A = PDP^{-1}$ y

$$\begin{aligned} A^k = PD^kP^{-1} &= \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3^k & 0 \\ 0 & 2^k \end{bmatrix}\begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3^k & 0 \\ 0 & 2^k \end{bmatrix}\begin{bmatrix} 1 & -2 \\ -1 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3^{k+1} - 2^{k+1} & -2(3^{k+1}) + 3(2^{k+1}) \\ 3^k - 2^k & -2(3^k) + 3(2^k) \end{bmatrix} \end{aligned}$$

Ahora se tiene que

$$\begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} = \mathbf{x}_n = A^{n-2}\mathbf{x}_2 = \begin{bmatrix} 3^{n-1} - 2^{n-1} & -2(3^{n-1}) + 3(2^{n-1}) \\ 3^{n-2} - 2^{n-2} & -2(3^{n-2}) + 3(2^{n-2}) \end{bmatrix}\begin{bmatrix} 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 3^n - 2^n \\ 3^{n-1} - 2^{n-1} \end{bmatrix}$$


de la cual se lee las solución $x_n = 3^n - 2^n$. (Para comprobar el trabajo, podría sustituir $n = 1, 2, \ldots, 5$ para verificar que esta fórmula produce los mismos términos que se calcularon usando la relación de recurrencia. ¡Inténtelo!)

Observe que x_n es una combinación lineal de potencias de los eigenvalores. Éste es necesariamente el caso en tanto los eigenvalores sean distintos [como lo hará explícito el Teorema 4.38(a)]. Con esta observación puede ahorrarse algo de trabajo. Una vez que calcule los eigenvalores $\lambda_1 = 3$ y $\lambda_2 = 2$, inmediatamente puede escribir

$$x_n = c_1 3^n + c_2 2^n$$

donde deben determinarse c_1 y c_2. Con las condiciones iniciales se tiene

$$1 = x_1 = c_1 3^1 + c_2 2^1 = 3c_1 + 2c_2$$

cuando $n = 1$ y

$$5 = x_2 = c_1 3^2 + c_2 2^2 = 9c_1 + 4c_2$$

cuando $n = 2$. Ahora resuelva el sistema

$$\begin{aligned} 3c_1 + 2c_2 &= 1 \\ 9c_1 + 4c_2 &= 5 \end{aligned}$$

para c_1 y c_2 para obtener $c_1 = 1$ y $c_2 = -1$. Por tanto, $x_n = 3^n - 2^n$, como antes.

Éste es el método que se usará en la práctica. Ahora se ilustrará su uso para encontrar una fórmula explícita para los números de Fibonacci.

Ejemplo 4.41 Resuelva la recurrencia de Fibonacci $f_0 = 0, f_1 = 1$ y $f_n = f_{n-1} + f_{n-2}$ para $n \geq 2$.

Solución Al escribir la recurrencia como $f_n - f_{n-1} - f_{n-2} = 0$, se ve que la ecuación característica es $\lambda^2 - \lambda - 1 = 0$, de modo que los eigenvalores son

$$\lambda_1 = \frac{1 + \sqrt{5}}{2} \qquad \text{y} \qquad \lambda_2 = \frac{1 - \sqrt{5}}{2}$$

De la discusión anterior se tiene que la solución a la relación de recurrencia tiene la forma

$$f_n = c_1\lambda_1^n + c_2\lambda_2^n = c_1\left(\frac{1 + \sqrt{5}}{2}\right)^n + c_2\left(\frac{1 - \sqrt{5}}{2}\right)^n$$

para algunos escalares c_1 y c_2.

Al usar las condiciones iniciales se encuentra

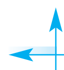
$$0 = f_0 = c_1\lambda_1^0 + c_2\lambda_2^0 = c_1 + c_2$$

y

$$1 = f_1 = c_1\lambda_1^1 + c_2\lambda_2^1 = c_1\left(\frac{1 + \sqrt{5}}{2}\right) + c_2\left(\frac{1 - \sqrt{5}}{2}\right)$$

Al resolver para c_1 y c_2, se obtiene $c_1 = 1/\sqrt{5}$ y $c_2 = -1/\sqrt{5}$. Por tanto, una fórmula explícita para el n-ésimo número de Fibonacci es

$$f_n = \frac{1}{\sqrt{5}}\left(\frac{1 + \sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1 - \sqrt{5}}{2}\right)^n \tag{5}$$



Jacques Binet (1786-1856) hizo aportaciones a la teoría de matrices, la teoría de números, física y astronomía. Descubrió la regla para la multiplicación de matrices en 1812. La fórmula de Binet para los números de Fibonacci en realidad se debe a Euler, quien la publicó en 1765; sin embargo, quedó en el olvido hasta que Binet publicó su versión en 1843. Como Cauchy, Binet fue realista y perdió su posición en la universidad cuando Carlos X abdicó en 1830. Recibió muchos honores por su trabajo, incluida su elección, en 1843, a la Académie des Sciences.

La fórmula (5) es una fórmula notable, porque se define en términos del número *irracional* $\sqrt{5}$, ¡aunque los números de Fibonacci son todos enteros! Trate de sustituir algunos valores para n para ver cómo se cancelan los términos $\sqrt{5}$ para dejar los valores enteros f_n. La fórmula (5) se conoce como ***fórmula de Binet***.

El método que acaba de bosquejarse funciona para cualquier relación de recurrencia lineal de segundo orden cuyos eigenvalores asociados sean todos distintos. Cuando haya un eigenvalor repetido, la técnica deberá modificarse, pues el método de diagonalización empleado puede ya no funcionar. El siguiente teorema resume ambas situaciones.

Teorema 4.38 Sea $x_n = ax_{n-1} + bx_{n-2}$ una relación de recurrencia que se define mediante una sucesión (x_n). Sean λ_1 y λ_2 los eigenvalores de la ecuación característica asociada $\lambda^2 - a\lambda - b = 0$.

a. Si $\lambda_1 \neq \lambda_2$; entonces $x_n = c_1\lambda_1^n + c_1\lambda_2^n$ para algunos escalares c_1 y c_2.
b. Si $\lambda_1 = \lambda_2 = \lambda$, entonces $x_n = c_1\lambda^n + c_2n\lambda^n$ para algunos escalares c_1 y c_2.

En cualquier caso, c_1 y c_2 pueden determinarse usando las condiciones iniciales.

Demostración (a) Al generalizar la discusión anterior, puede escribir la recurrencia como $\mathbf{x}_n = A\mathbf{x}_{n-1}$, donde

$$\mathbf{x}_n = \begin{bmatrix} x_n \\ x_{n-1} \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}$$

Dado que A tiene eigenvalores distintos, puede diagonalizarse. El resto de los detalles se dejan para el ejercicio 53.

(b) Se demostrará que $x_n = c_1\lambda^n + c_2n\lambda^n$ satisface la relación de recurrencia $x_n = ax_{n-1} + bx_{n-2}$ o, de manera equivalente,

$$x_n - ax_{n-1} - bx_{n-2} = 0 \tag{6}$$

si $\lambda^2 - a\lambda - b = 0$. Dado que

$$x_{n-1} = c_1\lambda^{n-1} + c_2(n-1)\lambda^{n-1} \quad \text{y} \quad x_{n-2} = c_1\lambda^{n-2} + c_2(n-2)\lambda^{n-2}$$

la sustitución en la ecuación (6) produce

$$\begin{aligned} x_n - ax_{n-1} - bx_{n-2} &= (c_1\lambda^n + c_2n\lambda^n) - a(c_1\lambda^{n-1} + c_2(n-1)\lambda^{n-1}) \\ &\quad - b(c_1\lambda^{n-2} + c_2(n-2)\lambda^{n-2}) \\ &= c_1(\lambda^n - a\lambda^{n-1} - b\lambda^{n-2}) + c_2(n\lambda^n - a(n-1)\lambda^{n-1} \\ &\quad - b(n-2)\lambda^{n-2}) \\ &= c_1\lambda^{n-2}(\lambda^2 - a\lambda - b) + c_2n\lambda^{n-2}(\lambda^2 - a\lambda - b) + c_2\lambda^{n-2}(a\lambda + 2b) \\ &= c_1\lambda^{n-2}(0) + c_2n\lambda^{n-2}(0) + c_2\lambda^{n-2}(a\lambda + 2b) \\ &= c_2\lambda^{n-2}(a\lambda + 2b) \end{aligned}$$

Pero, puesto que λ es una raíz doble de $\lambda^2 - a\lambda - b = 0$, se debe tener $a^2 + 4b = 0$ y $\lambda = a/2$, usando la fórmula cuadrática. En consecuencia, $a\lambda + 2b = a^2/2 + 2b = -4b/2 + 2b = 0$, de modo que

$$x_n - ax_{n-1} - bx_{n-2} = c_2\lambda^{n-2}(a\lambda + 2b) = c_2\lambda^{n-2}(0) = 0$$

Suponga que las condiciones iniciales son $x_0 = r$ y $x_1 = s$. Entonces, en (a) o en (b) hay una solución única para c_1 y c_2. (Vea el ejercicio 54.)

Ejemplo 4.42 Resuelva la relación de recurrencia $x_0 = 1$, $x_1 = 6$ y $x_n = 6x_{n-1} - 9x_{n-2}$ para $n \geq 2$.

Solución La ecuación característica es $\lambda^2 - 6\lambda + 9 = 0$, que tiene $\lambda = 3$ como raíz doble. Por el Teorema 4.38(b), debe tener $x_n = c_1 3^n + c_2 n 3^n = (c_1 + c_2 n)3^n$. Dado que $1 = x_0 = c_1$ y $6 = x_1 = (c_1 + c_2)3$, se encuentra que $c_2 = 1$, de modo que

$$x_n = (1 + n)3^n$$

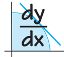
Las técnicas que se destacan en el Teorema 4.38 pueden extenderse a relaciones de recurrencia de orden superior. Se enuncia, sin demostración, el resultado general.

Teorema 4.39 Sea $x_n = a_{m-1}x_{n-1} + a_{m-2}x_{n-2} + \cdots + a_0 x_{n-m}$ una relación de recurrencia de orden m que se satisface mediante una sucesión (x_n). Suponga que el polinomio característico asociado

$$\lambda^m - a_{m-1}\lambda^{m-1} - a_{m-2}\lambda^{m-2} - \cdots - a_0$$

se factoriza como $(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2}\cdots(\lambda - \lambda_k)^{m_k}$, donde $m_1 + m_2 + \cdots + m_k = m$. Entonces x_n tiene la forma

$$x_n = (c_{11}\lambda_1^n + c_{12}n\lambda_1^n + c_{13}n^2\lambda_1^n + \cdots + c_{1m_1}n^{m_1-1}\lambda_1^n) + \cdots$$
$$+ (c_{k1}\lambda_k^n + c_{k2}n\lambda_k^n + c_{k3}n^2\lambda_k^n + \cdots + c_{km_k}n^{m_k-1}\lambda_k^n)$$



Sistemas de ecuaciones diferenciales lineales

En cálculo se aprende que si $x = x(t)$ es una función derivable que satisface una ecuación diferencial de la forma $x' = kx$, donde k es una constante, entonces la solución general es $x = Ce^{kt}$, donde C es una constante. Si se especifica una condición inicial $x(0) = x_0$, entonces, al sustituir $t = 0$ en la solución general, se encuentra que $C = x_0$. En consecuencia, la solución única a la ecuación diferencial que satisface la condición inicial es

$$x = x_0 e^{kt}$$

Suponga que tiene n funciones derivables de t, por decir, $x_1, x_2, \ldots, x_n$, que satisfacen un ***sistema de ecuaciones diferenciales***

$$\begin{aligned} x_1' &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ x_2' &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ &\vdots \\ x_n' &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n \end{aligned}$$

Puede escribir este sistema en forma matricial como $\mathbf{x}' = A\mathbf{x}$, donde

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad \mathbf{x}'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \quad \text{y} \quad A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

Ahora puede usar métodos matriciales para encontrar la solución.

Primero, una observación útil. Suponga que quiere resolver el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned} x_1' &= 2x_1 \\ x_2' &= 5x_2 \end{aligned}$$

Cada ecuación puede resolverse por separado, como anteriormente, para dar como resultado

$$\begin{aligned} x_1 &= C_1e^{2t} \\ x_2 &= C_2e^{5t} \end{aligned}$$

donde C_1 y C_2 con constantes. Note que, en forma matricial, la ecuación $\mathbf{x}' = A\mathbf{x}$ tiene una matriz de coeficientes *diagonal*

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 5 \end{bmatrix}$$

y los eigenvalores 2 y 5 ocurren en los exponenciales e^{2t} y e^{5t} de la solución. Esto sugiere que, para un sistema arbitrario, debe comenzar por diagonalizar la matriz de coeficientes si es posible.

Ejemplo 4.43

Resuelva el siguiente sistema de ecuaciones diferenciales:

$$\begin{aligned} x_1' &= x_1 + 2x_2 \\ x_2' &= 3x_1 + 2x_2 \end{aligned}$$

Solución Aquí la matriz de coeficientes es $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$, y se encuentra que los eigenvalores son $\lambda_1 = 4$ y $\lambda_2 = -1$, con sus correspondientes eigenvectores $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, respectivamente. Por tanto, A es diagonalizable y la matriz P que hace el trabajo es

$$P = [\mathbf{v}_1 \quad \mathbf{v}_2] = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$$

Se sabe que

$$P^{-1}AP = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix} = D$$

Sea $\mathbf{x} = P\mathbf{y}$ (de modo que $\mathbf{x}' = P\mathbf{y}'$) y sustituya estos resultados en la ecuación original $\mathbf{x}' = A\mathbf{x}$ para obtener $P\mathbf{y}' = AP\mathbf{y}$ o, de manera equivalente,

$$\mathbf{y}' = P^{-1}AP\mathbf{y} = D\mathbf{y}$$

Este es justo el sistema

$$\begin{aligned} y_1' &= 4y_1 \\ y_2' &= -y_2 \end{aligned}$$

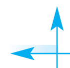
cuya solución general es

$$\begin{aligned} y_1 &= C_1e^{4t} \\ y_2 &= C_2e^{-t} \end{aligned} \quad \text{o} \quad \mathbf{y} = \begin{bmatrix} C_1e^{4t} \\ C_2e^{-t} \end{bmatrix}$$

Para encontrar **x**, sólo calcule

$$\mathbf{x} = P\mathbf{y} = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} C_1e^{4t} \\ C_2e^{-t} \end{bmatrix} = \begin{bmatrix} 2C_1e^{4t} - C_2e^{-t} \\ 3C_1e^{4t} + C_2e^{-t} \end{bmatrix}$$


de modo que $x_1 = 2C_1e^{4t} - C_2e^{-t}$ y $x_2 = 3C_1e^{4t} + C_2e^{-t}$. (Compruebe que estos valores satisfacen el sistema dado.)



Comentario Observe que la solución del ejemplo 4.43 también podría expresarse como

$$\mathbf{x} = C_1e^{4t}\begin{bmatrix} 2 \\ 3 \end{bmatrix} + C_2e^{-t}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = C_1e^{4t}\mathbf{v}_1 + C_2e^{-t}\mathbf{v}_2$$

Esta técnica se generaliza fácilmente a sistemas de $n \times n$ donde la matriz de coeficientes es diagonalizable. El siguiente teorema, cuya demostración se deja como ejercicio, resume la situación.

Teorema 4.40 Sea A una matriz diagonalizable de $n \times n$ y sea $P = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n]$ tal que

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Entonces la solución general al sistema $\mathbf{x}' = A\mathbf{x}$ es

$$\mathbf{x} = C_1e^{\lambda_1 t}\mathbf{v}_1 + C_2e^{\lambda_2 t}\mathbf{v}_2 + \cdots + C_ne^{\lambda_n t}\mathbf{v}_n$$

El siguiente ejemplo involucra un modelo biológico en el que dos especies viven en el mismo ecosistema. Es razonable suponer que la tasa de crecimiento de cada especie depende de los tamaños de *ambas* poblaciones. (Desde luego, existen otros factores que gobiernan el crecimiento, pero el modelo se mantendrá simple al ignorarlos.)

Si $x_1(t)$ y $x_2(t)$ denotan los tamaños de las dos poblaciones en el tiempo t, entonces $x_1'(t)$ y $x_2'(t)$ son sus tasas de crecimiento en el tiempo t. El modelo es de la forma

$$\begin{aligned} x_1'(t) &= ax_1(t) + bx_2(t) \\ x_2'(t) &= cx_1(t) + dx_2(t) \end{aligned}$$

donde los coeficientes a, b, c y d dependen de las condiciones.

Ejemplo 4.44 Mapaches y ardillas habitan el mismo ecosistema y compiten mutuamente por alimento, agua y espacio. Sean $r(t)$ y $s(t)$ las poblaciones de mapaches y ardillas en el tiempo t años, respectivamente. En ausencia de ardillas, la tasa de crecimiento de mapaches es $r'(t) = 2.5r(t)$, pero cuando están presentes las ardillas, la competencia frena la tasa de crecimiento de los mapaches a $r'(t) = 2.5r(t) - s(t)$. La población de ardillas es similarmente afectada por los mapaches. En ausencia de mapaches, la tasa de crecimiento de la población de ardillas es $s'(t) = 2.5s(t)$ y la tasa de crecimiento poblacional para las ardillas cuando comparten el ecosistema con los mapaches es $s'(t) = -0.25r(t) + 2.5s(t)$. Su-

ponga que inicialmente hay 60 mapaches y 60 ardillas en el ecosistema. Determine lo que ocurre con estas dos poblaciones.

Solución El sistema es $\mathbf{x}' = A\mathbf{x}$, donde

$$\mathbf{x} = \mathbf{x}(t) = \begin{bmatrix} r(t) \\ s(t) \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 2.5 & -1.0 \\ -0.25 & 2.5 \end{bmatrix}$$

Los eigenvalores de A son $\lambda_1 = 3$ y $\lambda_2 = 2$, con sus correspondientes eigenvectores $\mathbf{v}_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Por el Teorema 4.40, la solución general del sistema es

$$\mathbf{x}(t) = C_1e^{3t}\mathbf{v}_1 + C_2e^{2t}\mathbf{v}_2 = C_1e^{3t}\begin{bmatrix} -2 \\ 1 \end{bmatrix} + C_2e^{2t}\begin{bmatrix} 2 \\ 1 \end{bmatrix} \tag{7}$$

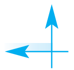
El vector de población inicial es $\mathbf{x}(0) = \begin{bmatrix} r(0) \\ s(0) \end{bmatrix} = \begin{bmatrix} 60 \\ 60 \end{bmatrix}$, de modo que, al hacer $t = 0$ en la ecuación (7), se tiene

$$C_1\begin{bmatrix} -2 \\ 1 \end{bmatrix} + C_2\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 60 \\ 60 \end{bmatrix}$$

Al resolver esta ecuación, se encuentra $C_1 = 15$ y $C_2 = 45$. Por tanto,

$$\mathbf{x}(t) = 15e^{3t}\begin{bmatrix} -2 \\ 1 \end{bmatrix} + 45e^{2t}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

de donde se encuentra $r(t) = -30e^{3t} + 90e^{2t}$ y $s(t) = 15e^{3t} + 45e^{2t}$. La figura 4.20 muestra las gráficas de estas dos funciones, y puede ver claramente que la población de mapaches muere después de poco más de un año.  (¿Puede determinar *exactamente* cuándo muere?)



Ahora considere un ejemplo similar, en el que una especie es una fuente de alimento para la otra. Tal modelo se llama ***modelo depredador-presa***. Una vez más, el modelo se simplificará drásticamente con la finalidad de ilustrar sus principales características.

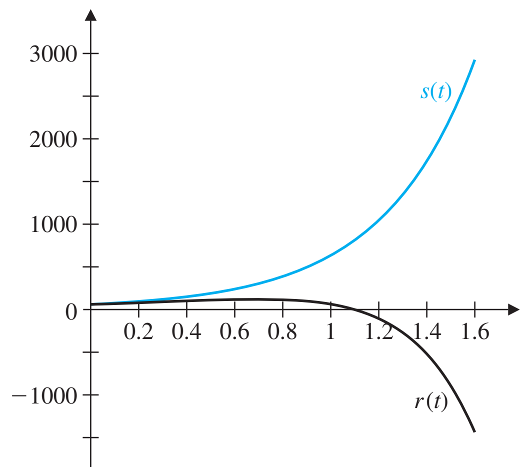


Figura 4.20
Poblaciones de mapaches y ardillas

a + bi **Ejemplo 4.45** Petirrojos y gusanos cohabitan en un ecosistema. Los petirrojos comen gusanos, que son su única fuente de alimento. Las poblaciones de petirrojos y gusanos en el tiempo t años se denotan mediante $r(t)$ y $w(t)$, respectivamente, y las ecuaciones que gobiernan el crecimiento de las dos poblaciones son

$$\begin{aligned} r'(t) &= w(t) - 12 \\ w'(t) &= -r(t) + 10 \end{aligned} \tag{8}$$

Si inicialmente 6 petirrojos y 20 gusanos ocupan el ecosistema, determine el comportamiento de las dos poblaciones con el tiempo.

Solución Lo primero que se nota acerca de este ejemplo es la presencia de las constantes adicionales, -12 y 10, en las dos ecuaciones. Por fortuna, puede deshacerse de ellas con un simple cambio de variables. Si hace $r(t) = x(t) + 10$ y $w(t) = y(t) + 12$, entonces $r'(t) = x'(t)$ and $w'(t) = y'(t)$. Al sustituir en las ecuaciones (8), se tiene

$$\begin{aligned} x'(t) &= y(t) \\ y'(t) &= -x(t) \end{aligned} \tag{9}$$

con las que es más fácil trabajar. Las ecuaciones (9) tienen la forma $\mathbf{x}' = A\mathbf{x}$, donde $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Las nuevas condiciones iniciales son

$$x(0) = r(0) - 10 = 6 - 10 = -4 \qquad \text{y} \qquad y(0) = w(0) - 12 = 20 - 12 = 8$$

de modo que $\mathbf{x}(0) = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$.

Al proceder como en el último ejemplo, se encuentran los eigenvalores y eigenvectores de A. El polinomio característico es $\lambda^2 + 1$, que no tiene raíces reales. ¿Qué debe hacer? No se tiene otra opción que usar las raíces complejas, que son $\lambda_1 = i$ y $\lambda_2 = -i$. Los correspondientes eigenvectores también son complejos, a saber, $\mathbf{v}_1 = \begin{bmatrix} 1 \\ i \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} 1 \\ -i \end{bmatrix}$. Por el Teorema 4.40, la solución tiene la forma

$$\mathbf{x}(t) = C_1 e^{it}\mathbf{v}_1 + C_2 e^{-it}\mathbf{v}_2 = C_1 e^{it}\begin{bmatrix} 1 \\ i \end{bmatrix} + C_2 e^{-it}\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

A partir de $\mathbf{x}(0) = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$, se tiene

$$C_1\begin{bmatrix} 1 \\ i \end{bmatrix} + C_2\begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} -4 \\ 8 \end{bmatrix}$$

cuya solución es $C_1 = -2 - 4i$ y $C_2 = -2 + 4i$. De modo que la solución al sistema (9) es

$$\mathbf{x}(t) = (-2 - 4i)e^{it}\begin{bmatrix} 1 \\ i \end{bmatrix} + (-2 + 4i)e^{-it}\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

¿Qué se hará con esta solución? Petirrojos y gusanos habitan un mundo real, ¡aunque la solución involucra números complejos! Proceda intrépidamente y aplique la fórmula de Euler

$$e^{it} = \cos t + i \operatorname{sen} t$$



(Apéndice C) para obtener $e^{-it} = \cos(-t) + i \operatorname{sen}(-t) = \cos t - i \operatorname{sen} t$. Al sustituir se tiene

$$\begin{aligned}
\mathbf{x}(t) &= (-2 - 4i)(\cos t + i \operatorname{sen} t)\begin{bmatrix} 1 \\ i \end{bmatrix} + (-2 + 4i)(\cos t - i \operatorname{sen} t)\begin{bmatrix} 1 \\ -i \end{bmatrix} \\
&= \begin{bmatrix} (-2\cos t + 4 \operatorname{sen} t) + i(-4\cos t - 2 \operatorname{sen} t) \\ (4\cos t + 2 \operatorname{sen} t) + i(-2\cos t + 4 \operatorname{sen} t) \end{bmatrix} \\
&\quad + \begin{bmatrix} (-2\cos t + 4 \operatorname{sen} t) + i(4\cos t + 2 \operatorname{sen} t) \\ (4\cos t + 2 \operatorname{sen} t) + i(2\cos t - 4 \operatorname{sen} t) \end{bmatrix} \\
&= \begin{bmatrix} -4\cos t + 8 \operatorname{sen} t \\ 8\cos t + 4 \operatorname{sen} t \end{bmatrix}
\end{aligned}$$

Esto produce $x(t) = -4\cos t + 8 \operatorname{sen} t$ y $y(t) = 8\cos t + 4 \operatorname{sen} t$. Al poner todo en términos de las variables originales, se concluye que

$$r(t) = x(t) + 10 = -4\cos t + 8 \operatorname{sen} t + 10$$

y

$$w(t) = y(t) + 12 = 8\cos t + 4 \operatorname{sen} t + 12$$

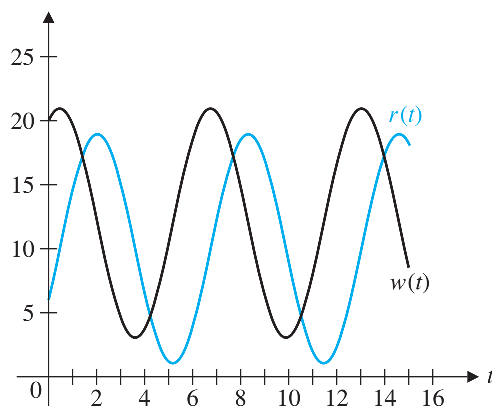


Figura 4.21

Poblaciones de petirrojos y gusanos

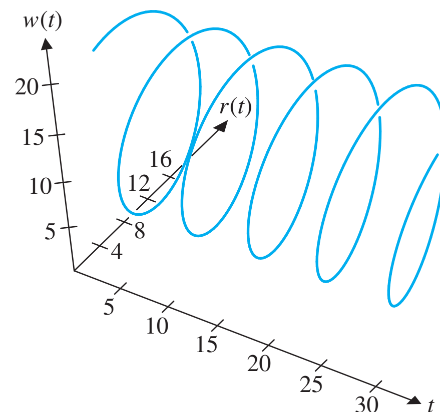
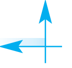


Figura 4.22

¡De modo que la solución es real después de todo! Las gráficas de $r(t)$ y $w(t)$ en la figura 4.21 muestran que las dos poblaciones oscilan periódicamente. Conforme aumenta la población de petirrojos, la población de gusanos comienza a disminuir, pero como disminuye la única fuente de alimento de los petirrojos, su número también comienza a declinar. A medida que los depredadores desaparecen, la población de gusanos comienza a recuperarse. Conforme aumenta su suministro de alimento, también lo hace la población de petirrojos y el ciclo se repite. Esta oscilación es típica de los ejemplos en los que los eigenvalores son complejos.

Graficar petirrojos, gusanos y tiempo en ejes separados, como en la figura 4.22, claramente revela la naturaleza cíclica de las dos poblaciones.



Esta sección concluye al observar lo que se hizo desde un punto de vista diferente. Si $x = x(t)$ es una función derivable de t, entonces la solución general de la ecuación diferencial ordinaria $x' = ax$ es $x = ce^{at}$, donde c es un escalar. Los sistemas de ecuaciones diferenciales lineales consideradas tienen la forma $\mathbf{x}' = A\mathbf{x}$, de modo que, si simplemente trabaja sin pensar, podría sentirse tentado a deducir que la solución sería $\mathbf{x} = \mathbf{c}e^{At}$, donde $\mathbf{c}$ es un vector. ¿Pero qué significa esto? En el lado derecho se tiene el *número e* elevado a la potencia de una *matriz*. Esto parece no tener sentido, pero verá que hay una forma de darle sentido.

Comience por considerar la expresión e^A. En cálculo, aprende que la función e^x tiene una expansión en serie de potencias

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

que converge para todo número real x. Por analogía, defina

$$e^A = I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots$$

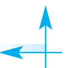
El lado derecho se define en términos de potencias de A, y puede demostrarse que converge para cualquier matriz real A. De modo que ahora e^A es una matriz, llamada ***exponencial*** de A. ¿Pero cómo puede calcularse e^A o e^{At}? Para matrices diagonales es sencillo.

Ejemplo 4.46

Calcule e^{Dt} para $D = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$.

Solución A partir de la definición se tiene

$$\begin{aligned} e^{Dt} &= I + Dt + \frac{(Dt)^2}{2!} + \frac{(Dt)^3}{3!} + \cdots \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 4t & 0 \\ 0 & -t \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} (4t)^2 & 0 \\ 0 & (-t)^2 \end{bmatrix} + \frac{1}{3!}\begin{bmatrix} (4t)^3 & 0 \\ 0 & (-t)^3 \end{bmatrix} + \cdots \\ &= \begin{bmatrix} 1 + (4t) + \frac{1}{2!}(4t)^2 + \frac{1}{3!}(4t)^3 + \cdots & 0 \\ 0 & 1 + (-t) + \frac{1}{2!}(-t)^2 + \frac{1}{3!}(-t)^3 + \cdots \end{bmatrix} \\ &= \begin{bmatrix} e^{4t} & 0 \\ 0 & e^{-t} \end{bmatrix} \end{aligned}$$



La exponencial de la matriz también es amigable si A es diagonalizable.

Ejemplo 4.47

Calcule e^A para $A = \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}$.

Solución En el ejemplo 4.43 se encontró que los eigenvalores de A son $\lambda_1 = 4$ y $\lambda_2 = -1$, con correspondientes eigenvectores $\mathbf{v}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, respectivamente. Por tanto, con $P = [\mathbf{v}_1 \quad \mathbf{v}_2] = \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}$, se tiene $P^{-1}AP = D = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$. Dado que $A = PDP^{-1}$, se tiene $A^k = PD^kP^{-1}$, de modo que

$$\begin{aligned} e^A &= I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots \\ &= PIP^{-1} + PDP^{-1} + \frac{1}{2!}PD^2P^{-1} + \frac{1}{3!}PD^3P^{-1} + \cdots \\ &= P\left(I + D + \frac{D^2}{2!} + \frac{D^3}{3!} + \cdots\right)P^{-1} \\ &= Pe^DP^{-1} \\ &= \begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} e^4 & 0 \\ 0 & e^{-1} \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 3 & 1 \end{bmatrix}^{-1} \\ &= \frac{1}{5}\begin{bmatrix} 2e^4 + 3e^{-1} & 2e^4 - 2e^{-1} \\ 3e^4 - 3e^{-1} & 3e^4 + 2e^{-1} \end{bmatrix} \end{aligned}$$

¡Ahora está en posición de demostrar que la audaz (y aparentemente absurda) suposición de una solución "exponencial" de $\mathbf{x}' = A\mathbf{x}$ no era tan disparatada del todo!

Teorema 4.41

Sea A una matriz diagonalizable de $n \times n$ con eigenvalores $\lambda_1, \lambda_2, \ldots, \lambda_n$. Entonces la solución general al sistema $\mathbf{x}' = A\mathbf{x}$ es $\mathbf{x} = e^{At}\mathbf{c}$, donde $\mathbf{c}$ es un vector constante arbitrario. Si se especifica una condición inicial $\mathbf{x}(0)$, entonces $\mathbf{c} = \mathbf{x}(0)$.

Demostración Sea P que diagonaliza A. Entonces $A = PDP^{-1}$ y, como en el ejemplo 4.47,

$$e^{At} = Pe^{Dt}P^{-1}$$

En consecuencia, es necesario comprobar que $\mathbf{x}' = A\mathbf{x}$ se satisface por $\mathbf{x} = Pe^{Dt}P^{-1}\mathbf{c}$. Ahora todo es constante, excepto por e^{Dt}, de modo que

$$\mathbf{x}' = \frac{d\mathbf{x}}{dt} = \frac{d}{dt}(Pe^{Dt}P^{-1}\mathbf{c}) = P\frac{d}{dt}(e^{Dt})P^{-1}\mathbf{c} \qquad (10)$$

Si
$$D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

entonces
$$e^{Dt} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix}$$

Al efectuar derivadas se tiene

$$\begin{aligned} \frac{d}{dt}(e^{Dt}) &= \begin{bmatrix} \frac{d}{dt}(e^{\lambda_1 t}) & 0 & \cdots & 0 \\ 0 & \frac{d}{dt}(e^{\lambda_2 t}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{d}{dt}(e^{\lambda_n t}) \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & \lambda_2 e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n e^{\lambda_n t} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix} \\ &= De^{Dt} \end{aligned}$$

Al sustituir este resultado en la ecuación (10) se obtiene

$$\mathbf{x}' = PDe^{Dt}P^{-1}\mathbf{c} = PDP^{-1}Pe^{Dt}P^{-1}\mathbf{c} = (PDP^{-1})(Pe^{Dt}P^{-1})\mathbf{c} = Ae^{At}\mathbf{c} = A\mathbf{x}$$

como se requería.

La última expresión se obtiene fácilmente del hecho de que si $\mathbf{x} = \mathbf{x}(t) = e^{At}\mathbf{c}$, entonces

$$\mathbf{x}(0) = e^{A\cdot 0}\mathbf{c} = e^{O}\mathbf{c} = I\mathbf{c} = \mathbf{c}$$



pues $e^{O} = I$. (¿Por qué?)

Por ejemplo, vea *Linear Algebra* de S. H. Friedberg, A. J. Insel y L. E. Spence (Englewood Cliffs, NJ: Prentice-Hall, 1979).

De hecho, el Teorema 4.41 es verdadero incluso si A no es diagonalizable, pero no se demostrará esto. El cálculo de exponenciales de matrices para matrices no diagonalizables requiere la *forma normal de Jordan* de una matriz, un tema que puede encontrarse en textos más avanzados de álgebra lineal.

De manera ideal, este breve paréntesis sirvió para ilustrar el poder de las matemáticas para generalizar y el valor del pensamiento creativo. Las exponenciales de matrices evidencian ser herramientas muy importantes en muchas aplicaciones del álgebra lineal, tanto teórica como aplicada.

a + bi

Sistemas dinámicos lineales discretos

Este capítulo concluye como empezó, observando sistemas dinámicos. Las cadenas de Markov y el modelo de Leslie de crecimiento poblacional son ejemplos de ***sistemas dinámicos lineales discretos***. Cada uno puede describirse mediante una ecuación matricial de la forma

$$\mathbf{x}_{k+1} = A\mathbf{x}_k$$

donde el vector $\mathbf{x}_k$ registra el estado del sistema en el "tiempo" k y A es una matriz cuadrada. Como se vio, el comportamiento a largo plazo de dichos sistemas se relaciona con los eigenvalores y eigenvectores de la matriz A. El método de potencia explota la naturaleza iterativa de tales sistemas dinámicos para aproximar eigenvalores y eigenvectores, y el teorema de Perron-Frobenius brinda información especializada acerca del comportamiento a largo plazo de un sistema dinámico lineal discreto cuya matriz de coeficientes A es no negativa.

Cuando A es una matriz de 2×2, puede describir geométricamente la evolución de un sistema dinámico. La ecuación $\mathbf{x}_{k+1} = A\mathbf{x}_k$ es en realidad una colección infinita de ecuaciones. Al comenzar con un vector inicial $\mathbf{x}_0$, se tiene:

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 \\ \mathbf{x}_2 &= A\mathbf{x}_1 \\ \mathbf{x}_3 &= A\mathbf{x}_2 \\ &\vdots \end{aligned}$$

El conjunto $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots\}$ se llama ***trayectoria*** del sistema. (Para propósitos gráficos, cada vector en una trayectoria se identificará con su punta, de modo que pueda graficarse como un punto.) Note que $\mathbf{x}_k = A^k\mathbf{x}_0$.

Ejemplo 4.48

Sea $A = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.8 \end{bmatrix}$. Para el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$, grafique los primeros cinco puntos en las trayectorias con los siguientes vectores iniciales:

(a) $\mathbf{x}_0 = \begin{bmatrix} 5 \\ 0 \end{bmatrix}$ (b) $\mathbf{x}_0 = \begin{bmatrix} 0 \\ -5 \end{bmatrix}$ (c) $\mathbf{x}_0 = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ (d) $\mathbf{x}_0 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$

Solución (a) Calcule $\mathbf{x}_1 = A\mathbf{x}_0 = \begin{bmatrix} 2.5 \\ 0 \end{bmatrix}$, $\mathbf{x}_2 = A\mathbf{x}_1 = \begin{bmatrix} 1.25 \\ 0 \end{bmatrix}$, $\mathbf{x}_3 = A\mathbf{x}_2 = \begin{bmatrix} 0.625 \\ 0 \end{bmatrix}$, $\mathbf{x}_4 = A\mathbf{x}_3 = \begin{bmatrix} 0.3125 \\ 0 \end{bmatrix}$. Éstos se grafican en la figura 4.23 y los puntos se conectan para resaltar la trayectoria. Cálculos similares producen las trayectorias marcadas (b), (c) y (d) en la figura 4.23.

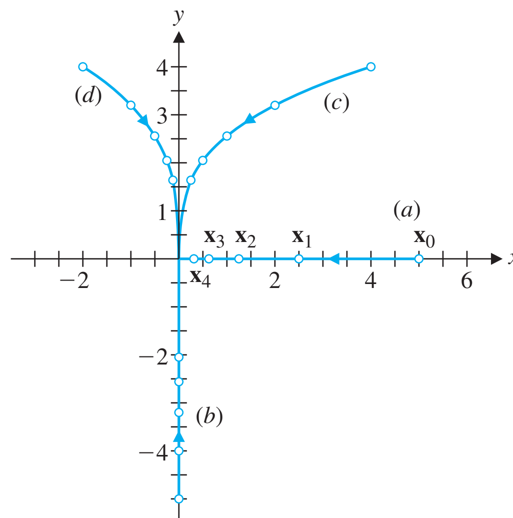


Figura 4.23

En el ejemplo 4.48 toda trayectoria converge en **0**. El origen se llama ***atractor*** en este caso. Puede entender por qué es así a partir del Teorema 4.19. La matriz A en el ejemplo 4.48 tiene eigenvectores $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ correspondientes a sus eigenvalores 0.5 y 0.8, respectivamente. (Compruebe esto.) En concordancia, para cualquier vector inicial

$$\mathbf{x}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = c_1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

se tiene

$$\mathbf{x}_k = A^k\mathbf{x}_0 = c_1(0.5)^k\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(0.8)^k\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Puesto que $(0.5)^k$ y $(0.8)^k$ tienden a cero conforme k se vuelve grande, $\mathbf{x}_k$ tiende a **0** para cualquier elección de $\mathbf{x}_0$. Además, del Teorema 4.28 se sabe que, puesto que 0.8 es el eigenvalor dominante de A, $\mathbf{x}_k$ tenderá a un múltiplo del correspondiente eigenvector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ en tanto $c_2 \neq 0$ (el coeficiente de $\mathbf{x}_0$ correspondiente a $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$). En otras palabras, todas las trayectorias, excepto las que comienzan en el eje x (donde $c_2 = 0$) tenderán al eje y, como muestra la figura 4.23.

Ejemplo 4.49

Analice el comportamiento del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ correspondiente a la matriz $A = \begin{bmatrix} 0.65 & -0.15 \\ -0.15 & 0.65 \end{bmatrix}$.

Solución Los eigenvalores de A son 0.5 y 0.8, con correspondientes eigenvectores $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, respectivamente. (Compruebe esto.) Por tanto, para un vector inicial $\mathbf{x}_0 = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, se tiene

$$\mathbf{x}_k = A^k\mathbf{x}_0 = c_1(0.5)^k\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2(0.8)^k\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Una vez más, el origen es un atractor, porque $\mathbf{x}_k$ tiende a **0** para cualquier elección de $\mathbf{x}_0$. Si $c_2 \neq 0$, la trayectoria tenderá a la recta que pasa por el origen con vector director $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. En la figura 4.24 se muestran varias de tales trayectorias. Los vectores $\mathbf{x}_0$ donde $c_2 = 0$ están sobre la recta que pasa por el origen con vector director $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, y la trayectoria correspondiente en este caso sigue esta recta hacia el origen.

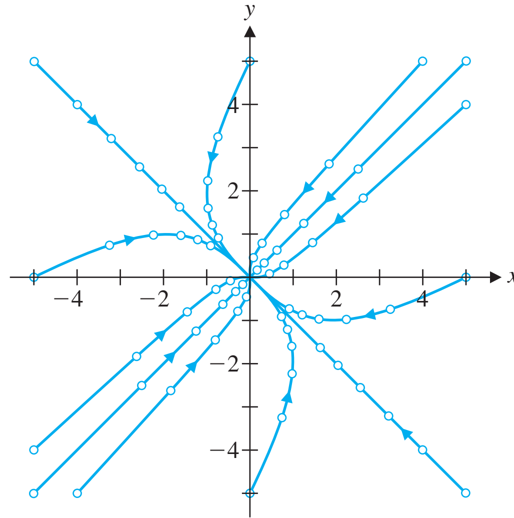


Figura 4.24

Ejemplo 4.50

Analice el comportamiento de los sistemas dinámicos $\mathbf{x}_{k+1} = A\mathbf{x}_k$ correspondientes a las siguientes matrices:

(a) $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$

Solución (a) Los eigenvalores de A son 5 y 3 con sus correspondientes eigenvectores $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, respectivamente. En consecuencia, para un vector inicial $\mathbf{x}_0 = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, se tiene

$$\mathbf{x}_k = A^k\mathbf{x}_0 = c_1 5^k\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 3^k\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Conforme k se vuelve grande, también lo hacen 5^k y 3^k. En consecuencia, $\mathbf{x}_k$ se aleja del origen. Puesto que el eigenvalor dominante de 5 tiene correspondiente eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, todas las trayectorias para las que $c_1 \neq 0$ eventualmente terminarán en el primero o tercer cuadrantes. Las trayectorias con $c_1 = 0$ comenzarán y permanecerán en la recta $y = -x$, cuyo vector director es $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Vea la figura 4.25(a).

(b) En este ejemplo, los eigenvalores son 1.5 y 0.5, con sus correspondientes eigenvectores $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ y $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$, respectivamente. Por tanto,

$$\mathbf{x}_k = c_1(1.5)^k\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2(0.5)^k\begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \text{si} \quad \mathbf{x}_0 = c_1\begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

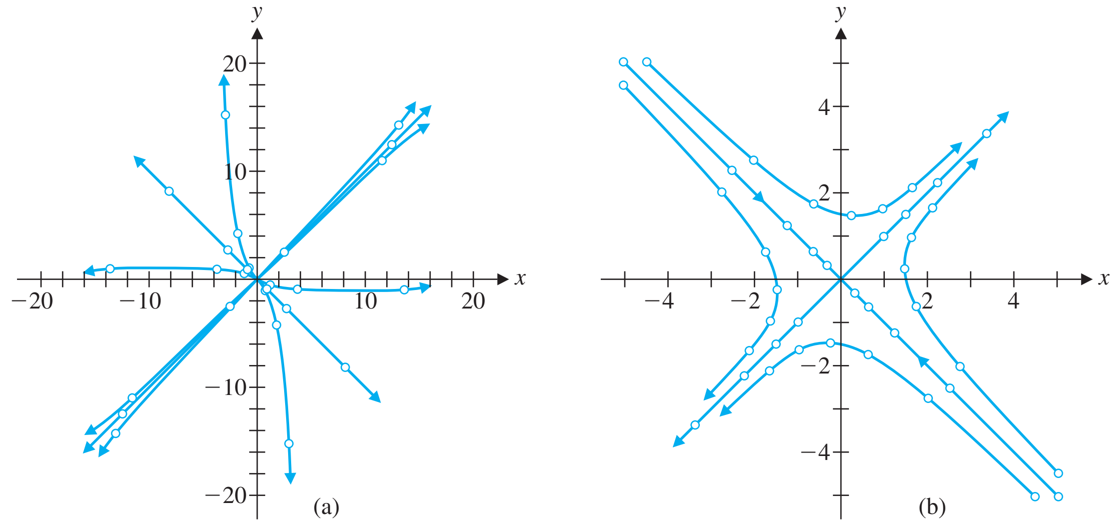


Figura 4.25

Si $c_1 = 0$, entonces $\mathbf{x}_k = c_2(0.5)^k\begin{bmatrix}-1\\1\end{bmatrix} \to \begin{bmatrix}0\\0\end{bmatrix}$ conforme $k \to \infty$. Pero si $c_1 \neq 0$, entonces

$$\mathbf{x}_k = c_1(1.5)^k\begin{bmatrix}1\\1\end{bmatrix} + c_2(0.5)^k\begin{bmatrix}-1\\1\end{bmatrix} \approx c_1(1.5)^k\begin{bmatrix}1\\1\end{bmatrix} \quad \text{conforme} \quad k \to \infty$$

y tales trayectorias se aproximan asintóticamente a la recta $y = x$. Vea la figura 4.25(b).

En el ejemplo 4.50(a), todos los puntos que comienzan cerca del origen se vuelven cada vez más grandes en magnitud, porque $|\lambda| > 1$ para ambos eigenvalores; **0** se llama ***repulsor***. En el ejemplo 4.50(b), **0** se llama ***punto de silla***, porque el origen atrae puntos en algunas direcciones y los repele en otras direcciones. En este caso, $|\lambda_1| < 1.$ y $|\lambda_2| > 1.$

El siguiente ejemplo muestra lo que puede suceder cuando los eigenvalores de una matriz real de 2×2 son complejos (y por tanto conjugado de otro).

Ejemplo 4.51

Grafique la trayectoria que comienza con $\mathbf{x}_0 = \begin{bmatrix}4\\4\end{bmatrix}$ para los sistemas dinámicos $\mathbf{x}_{k+1} = A\mathbf{x}_k$ que corresponden a las siguientes matrices:

(a) $A = \begin{bmatrix}0.5 & -0.5\\0.5 & 0.5\end{bmatrix}$ (b) $A = \begin{bmatrix}0.2 & -1.2\\0.6 & 1.4\end{bmatrix}$

Solución Las trayectorias se muestran en las figuras 4.26(a) y (b), respectivamente. Note que (a) es una trayectoria en espiral hacia el origen, mientras que (b) parece seguir una órbita elíptica.

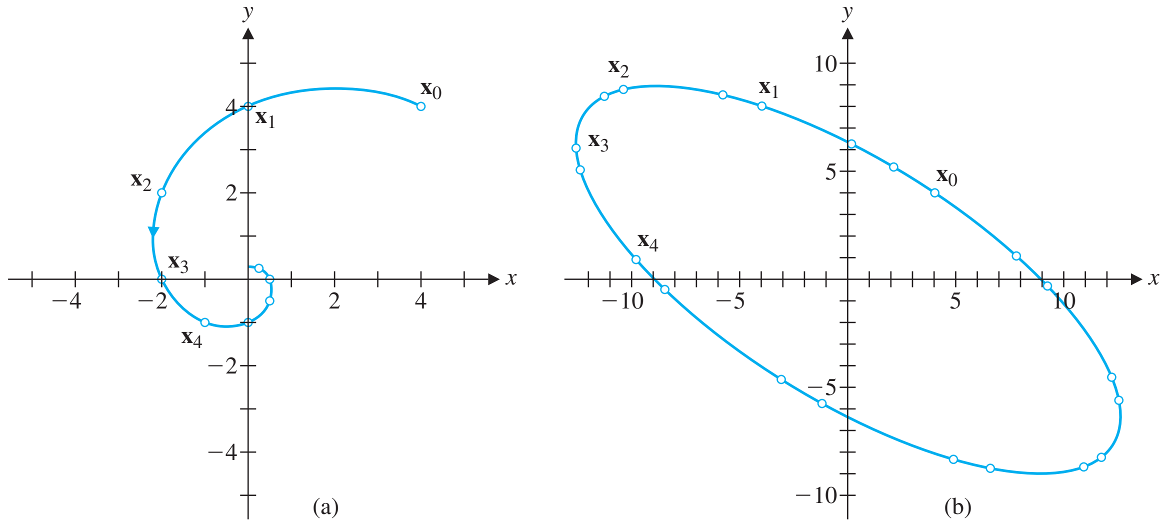


Figura 4.26

El siguiente teorema explica el comportamiento en espiral de la trayectoria del ejemplo 4.51(a).

Teorema 4.42 Sea $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. Los eigenvalores de A son $\lambda = a \pm bi$, y si a y b no son ambos cero, entonces A puede factorizarse como

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

donde $r = |\lambda| = \sqrt{a^2 + b^2}$ y θ es el argumento principal de $a + bi$.

Demostración Los eigenvalores de A son

$$\lambda = \tfrac{1}{2}(2a \pm \sqrt{4(-b^2)}) = \tfrac{1}{2}(2a \pm 2\sqrt{b^2}\sqrt{-1}) = a \pm |b|i = a \pm bi$$

por el ejercicio 35(b) de la sección 4.1. La figura 4.27 muestra $a + bi$, r y θ. Se tiene que

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = r\begin{bmatrix} a/r & -b/r \\ b\ r & a\ r \end{bmatrix} = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

Comentario Geométricamente, el Teorema 4.42 implica que cuando $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \neq O$, la transformación lineal $T(\mathbf{x}) = A\mathbf{x}$ es la composición de una rotación $R = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ a través del ángulo θ seguida por un escalamiento $S = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$ con factor r (figura 4.28). En el ejemplo 4.51(a), los eigenvalores son $\lambda = 0.5 \pm 0.5i$ de modo que $r = |\lambda| = \sqrt{2}/2 \approx 0.707 < 1$, y en consecuencia todas las trayectorias caen en espiral hacia $\mathbf{0}$.

El siguiente teorema muestra que, en general, cuando una matriz real de 2×2 tiene eigenvalores complejos, es similar a una matriz de la forma $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. Para un vector complejo

$$\mathbf{x} = \begin{bmatrix} z \\ w \end{bmatrix} = \begin{bmatrix} a + bi \\ c + di \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix}i$$

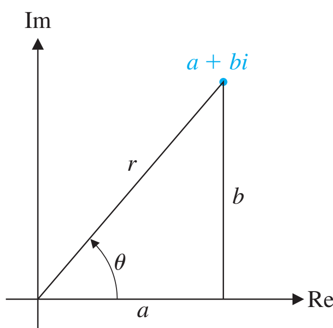


Figura 4.27

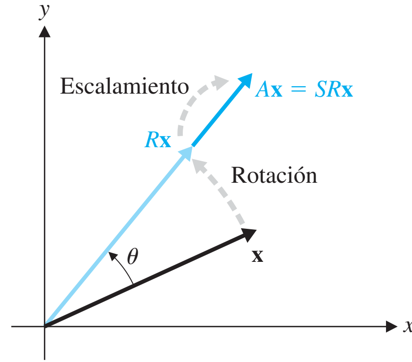


Figura 4.28
Rotación seguida por un escalamiento

defina la parte real, Re **x**, y la parte imaginaria, Im **x**, de **x** como

$$\text{Re}\,\mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \text{Re}z \\ \text{Re}w \end{bmatrix} \quad \text{y} \quad \text{Im}\,\mathbf{x} = \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} \text{Im}z \\ \text{Im}w \end{bmatrix}$$

Teorema 4.43

Sea A una matriz real de 2×2 con un eigenvalor complejo $\lambda = a - bi$ (donde $b \neq 0$ y correspondiente eigenvector $\mathbf{x}$. Entonces la matriz $P = [\text{Re}\,\mathbf{x} \quad \text{Im}\,\mathbf{x}]$ es invertible y

$$A = P\begin{bmatrix} a & -b \\ b & a \end{bmatrix}P^{-1}$$

Demostración Sea $\mathbf{x} = \mathbf{u} + \mathbf{v}i$, de modo que Re $\mathbf{x} = \mathbf{u}$ e Im $\mathbf{x} = \mathbf{v}$. De $A\mathbf{x} = \lambda\mathbf{x}$, se tiene

$$\begin{aligned} A\mathbf{u} + A\mathbf{v}i = A\mathbf{x} = \lambda\mathbf{x} &= (a - bi)(\mathbf{u} + \mathbf{v}i) \\ &= a\mathbf{u} + a\mathbf{v}i - b\mathbf{u}i + b\mathbf{v} = (a\mathbf{u} + b\mathbf{v}) + (-b\mathbf{u} + a\mathbf{v})i \end{aligned}$$

Al igualar las partes real e imaginaria se obtiene

$$A\mathbf{u} = a\mathbf{u} + b\mathbf{v} \quad \text{y} \quad A\mathbf{v} = -b\mathbf{u} + a\mathbf{v}$$

Ahora $P = [\mathbf{u} \mid \mathbf{v}]$, de modo que

$$\begin{aligned} P\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = [\mathbf{u} \mid \mathbf{v}]\begin{bmatrix} a & -b \\ b & a \end{bmatrix} &= [a\mathbf{u} + b\mathbf{v} \mid -b\mathbf{u} + a\mathbf{v}] = [A\mathbf{u} \mid A\mathbf{v}] = A[\mathbf{u} \mid \mathbf{v}] \\ &= AP \end{aligned}$$

Para demostrar que P es invertible, es suficiente con demostrar que **u** y **v** son linealmente independientes. Si **u** y **v** no son linealmente independientes, entonces se tendría que $\mathbf{v} = k\mathbf{u}$ para algún escalar k (complejo distinto de cero), porque ni **u** ni **v** son **0**. Por tanto,

$$\mathbf{x} = \mathbf{u} + \mathbf{v}i = \mathbf{u} + k\mathbf{u}i = (1 + ki)\mathbf{u}$$

Ahora, dado que A es real, $A\mathbf{x} = \lambda\mathbf{x}$ implica que

$$A\overline{\mathbf{x}} = \overline{A}\overline{\mathbf{x}} = \overline{A\mathbf{x}} = \overline{\lambda\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}$$

de modo que $\overline{\mathbf{x}} = \mathbf{u} - \mathbf{v}i$ es un eigenvector correspondiente al otro eigenvalor $\overline{\lambda} = a + bi$. Pero

$$\overline{\mathbf{x}} = \overline{(1 + ki)\mathbf{u}} = (1 - \overline{k}i)\mathbf{u}$$

dado que **u** es un vector real. En consecuencia, los eigenvectores **x** y $\overline{\mathbf{x}}$ de A son ambos múltiplos de **u** distintos de cero y por tanto son múltiplos mutuos. Esto es imposible porque los eigenvectores que corresponden a distintos eigenvalores deben ser linealmente independientes por el Teorema 4.20. (Este teorema es válido para los números complejos, así como para los números reales.)

Esta contradicción implica que **u** y **v** son linealmente independientes y en consecuencia P es invertible. Ahora se tiene que

$$A = P\begin{bmatrix} a & -b \\ b & a \end{bmatrix}P^{-1}$$

El Teorema 4.43 sirve para explicar el ejemplo 4.51(b). Los eigenvalores de $A = \begin{bmatrix} 0.2 & -1.2 \\ 0.6 & 1.4 \end{bmatrix}$ son $0.8 \pm 0.6i$. Para $\lambda = 0.8 - 0.6i$, un correspondiente eigenvector es

$$\mathbf{x} = \begin{bmatrix} -1-i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix}i$$

Del Teorema 4.43 se deduce que para $P = \begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix}$ y $C = \begin{bmatrix} 0.8 & -0.6 \\ 0.6 & 0.8 \end{bmatrix}$, se tiene

$$A = PCP^{-1} \quad \text{y} \quad P^{-1}AP = C$$

Para el sistema dinámico dado, $\mathbf{x}_{k+1} = A\mathbf{x}_k$, se realiza un cambio de variable. Sea

$$\mathbf{x}_k = P\mathbf{y}_k \quad \text{(o, de manera equivalente } \mathbf{y}_k = P^{-1}\mathbf{x}_k\text{)}$$

Entonces

$$P\mathbf{y}_{k+1} = \mathbf{x}_{k+1} = A\mathbf{x}_k = AP\mathbf{y}_k$$

de modo que

$$\mathbf{y}_{k+1} = \mathbf{x}_{k+1} = A\mathbf{x}_k = P^{-1}AP\mathbf{y}_k = C\mathbf{y}_k$$



Ahora C tiene los mismos eigenvalores que A (¿por qué?) y $|0.8 \pm 0.6i| = 1$. Por tanto, el sistema dinámico $\mathbf{y}_{k+1} = C\mathbf{y}_k$ simplemente rota los puntos en cada trayectoria en una circunferencia en torno al origen, por el Teorema 4.42.

Para determinar una trayectoria del sistema dinámico del ejemplo 4.51(b), se aplica iterativamente la transformación lineal $T(\mathbf{x}) = A\mathbf{x} = PCP^{-1}\mathbf{x}$. La transformación se puede considerar como la composición de un cambio de variable (de **x** a **y**), seguida por la rotación determinada por C, seguida por el cambio de variable inverso (**y** de vuelta a **x**). Esta idea se encontrará nuevamente en la sección 5.5 acerca de la aplicación a la graficación de ecuaciones cuadráticas y, de manera más general, como "cambio de base" en la sección 6.3. En el ejercicio 96 de la sección 5.5 se le mostrará que la trayectoria del ejemplo 4.51(b) de hecho es una elipse, como parece ser en la figura 4.26(b).

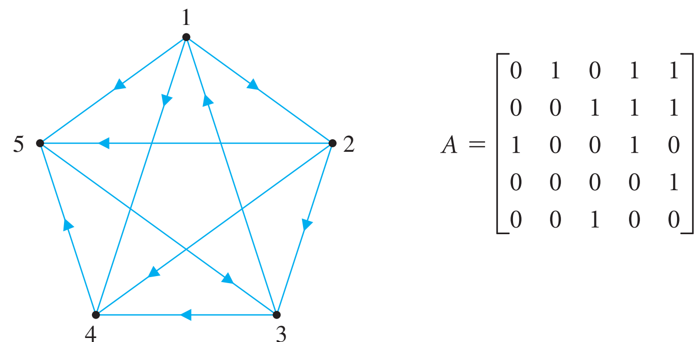
Para resumir: si una matriz A real de 2×2 tiene eigenvalores complejos $\lambda = a \pm bi$, entonces las trayectorias del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ giran en espiral hacia adentro si $|\lambda| < 1$ (**0** es un ***atractor espiral***), giran en espiral hacia afuera si $|\lambda| > 1$ (**0** es un ***repulsor espiral***) y se encuentran en una órbita cerrada si $|\lambda| = 1$ (**0** es un ***centro orbital***).

Clasificación de equipos deportivos y búsqueda en Internet

En cualquier liga deportiva de competencia, la clasificación de jugadores o equipos no necesariamente es un proceso directo. El mero conteo de triunfos y derrotas pasa por alto la posibilidad de que un equipo pueda acumular una gran cantidad de victorias contra equipos débiles, mientras que otro equipo puede tener menos victorias pero todas contra equipos fuertes. ¿Cuál de esos equipos es mejor? ¿Cómo deben compararse dos equipos que nunca jugaron entre sí? ¿Deben tomarse en cuenta los puntos anotados? ¿Los puntos en contra?

A pesar de estas complejidades, la clasificación de los atletas y los equipos deportivos se ha vuelto un lugar común y un aspecto muy esperado en los medios de comunicación. Por ejemplo, existen varias clasificaciones anuales de equipos estadounidenses de futbol americano colegial y basquetbol, y los golfistas y tenistas también se clasifican internacionalmente. Existen muchos esquemas con derechos de autor registrados que se usan para producir tales clasificaciones, pero puede obtener cierta comprensión de cómo abordar el problema usando las ideas de este capítulo.

Para establecer la idea básica, revise el ejemplo 3.69. Cinco tenistas juegan entre ellos en un torneo todos contra todos. Triunfos y derrotas se registran en forma de digrafo donde una arista dirigida de i a j indica que el jugador i venció al jugador j. Por ende, la correspondiente matriz de adyacencia A tiene $a_{ij} = 1$ si el jugador i vence al jugador j y tiene $a_{ij} = 0$ en caso contrario.



Sería apropiado asociar una clasificación r_i al jugador i de manera que $r_i > r_j$ indique que el jugador i está clasificado más alto que el jugador j. Con este propósito, se

requiere que las r_i sean probabilidades (esto es: $0 \leq r_i \leq 1$ para toda i, y $r_1 + r_2 + r_3 + r_4 + r_5 = 1$) y entonces organizar las clasificaciones en un *vector de clasificación*

$$\mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix}$$

Más aún, se insiste en que la clasificación del jugador i debe ser proporcional a la suma de las clasificaciones de los jugadores vencidos por el jugador i. Por ejemplo, el jugador 1 vence a los jugadores 2, 4 y 5, de modo que se quiere

$$r_1 = \alpha(r_2 + r_4 + r_5)$$

donde α es la constante de proporcionalidad. Al escribir ecuaciones similares para los otros jugadores se produce el siguiente sistema:

$$\begin{aligned} r_1 &= \alpha(r_2 + r_4 + r_5) \\ r_2 &= \alpha(r_3 + r_4 + r_5) \\ r_3 &= \alpha(r_1 + r_4) \\ r_4 &= \alpha r_5 \\ r_5 &= \alpha r_3 \end{aligned}$$

Observe que puede escribir este sistema en forma matricial como

$$\begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix} = \alpha \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \\ r_5 \end{bmatrix} \quad \text{o} \quad \mathbf{r} = \alpha A\mathbf{r}$$

De manera equivalente, vea que el vector de clasificación $\mathbf{r}$ debe satisfacer $A\mathbf{r} = \frac{1}{\alpha}\mathbf{r}$. En otras palabras, ¡$\mathbf{r}$ es un eigenvector que corresponde a la matriz A!

Más aún, A es una matriz no negativa primitiva, de modo que el Teorema de Perron-Frobenius garantiza la existencia de un vector de clasificación *único* $\mathbf{r}$. En este ejemplo, el vector de clasificación evidencia ser

$$\mathbf{r} = \begin{bmatrix} 0.29 \\ 0.27 \\ 0.22 \\ 0.08 \\ 0.14 \end{bmatrix}$$

de modo que los jugadores se clasificarían en el orden 1, 2, 3, 5, 4.

Al modificar la matriz A es posible tomar en cuenta muchas de las complejidades mencionadas en el párrafo de apertura. No obstante, este ejemplo simple sirvió para indicar un enfoque útil del problema de clasificar equipos.

La misma idea puede usarse para entender cómo funciona un buscador en Internet, como Google. Los buscadores más antiguos solían regresar los resultados de una búsqueda en *desorden*. Con frecuencia, los sitios útiles estaban enterrados entre los irrelevantes. Usualmente se requería mucho desplazamiento para descubrir lo que se buscaba. En contraste, Google regresa resultados de búsqueda *ordenados* de acuerdo con su probabilidad de relevancia. Por tanto, se necesita un método para clasificar sitios web.

En lugar de equipos que juegan entre sí, ahora se tienen sitios web vinculados mutuamente. Una vez más puede usar un digrafo para modelar la situación, sólo que ahora una arista de i a j indica que el sitio web i se vincula con (o se refiere a) el sitio web j. De este modo, mientras que para el digrafo de equipos deportivos las aristas dirigidas hacia adentro son malas (indican derrotas), para el digrafo de Internet las aristas dirigidas hacia adentro son buenas (indican vínculos desde otros sitios). En este escenario, se quiere que la clasificación del sitio web i sea proporcional a la suma de las clasificaciones de todos los sitios web que se vinculan con él.

Al usar el digrafo de la página 367 para representar sólo cinco sitios web, se tiene

$$r_4 = \alpha(r_1 + r_2 + r_3)$$

por ejemplo. Es fácil ver que ahora se quiere usar la *transpuesta* de la matriz de adyacencia del digrafo. Por tanto, el vector de clasificación $\mathbf{r}$ debe satisfacer $A^T\mathbf{r} = \dfrac{1}{\alpha}\mathbf{r}$ y por tanto será el eigenvector de Perron de A^T. En este ejemplo, se obtiene

$$A^T = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad \mathbf{r} = \begin{bmatrix} 0.14 \\ 0.08 \\ 0.22 \\ 0.27 \\ 0.29 \end{bmatrix}$$

de modo que una búsqueda que regrese estos cinco sitios los presentaría en el orden 5, 4, 3, 1, 2.

En realidad, Google utiliza una variante del método aquí descrito y calcula el vector de clasificación mediante un proceso iterativo muy similar al método de potencia (sección 4.5).

Ejercicios 4.6

Cadenas de Markov

¿Cuáles de las matrices estocásticas de los ejercicios 1-6 son regulares?

1. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

2. $\begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{1}{2} \end{bmatrix}$

3. $\begin{bmatrix} \frac{1}{3} & 1 \\ \frac{2}{3} & 0 \end{bmatrix}$

4. $\begin{bmatrix} \frac{1}{2} & 0 & 1 \\ \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

5. $\begin{bmatrix} 0.1 & 0 & 0.5 \\ 0.5 & 1 & 0 \\ 0.4 & 0 & 0.5 \end{bmatrix}$

6. $\begin{bmatrix} 0.5 & 1 & 0 \\ 0.5 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

En los ejercicios 7-9, P es la matriz de transición de una cadena de Markov regular. Encuentre la matriz de transición de rango largo L de P.

7. $P = \begin{bmatrix} \frac{1}{3} & \frac{1}{6} \\ \frac{2}{3} & \frac{5}{6} \end{bmatrix}$

8. $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{6} & \frac{1}{2} \end{bmatrix}$

9. $P = \begin{bmatrix} 0.2 & 0.3 & 0.4 \\ 0.6 & 0.1 & 0.4 \\ 0.2 & 0.6 & 0.2 \end{bmatrix}$

10. Demuestre que el vector de probabilidad de estado estacionario de una cadena de Markov regular es único. [*Sugerencia:* use los Teoremas 4.33 o 4.34.]

Crecimiento poblacional

En los ejercicios 11-14, calcule el eigenvalor positivo y su correspondiente eigenvector positivo de la matriz de Leslie L.

11. $L = \begin{bmatrix} 0 & 2 \\ 0.5 & 0 \end{bmatrix}$

12. $L = \begin{bmatrix} 1 & 1.5 \\ 0.5 & 0 \end{bmatrix}$

13. $L = \begin{bmatrix} 0 & 7 & 4 \\ 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \end{bmatrix}$

14. $L = \begin{bmatrix} 1 & 5 & 3 \\ \frac{1}{3} & 0 & 0 \\ 0 & \frac{2}{3} & 0 \end{bmatrix}$

15. Si una matriz de Leslie tiene un eigenvalor positivo único λ_1, ¿cuál es el significado para la población si $\lambda_1 > 1$? ¿Y si $\lambda_1 < 1$? ¿Y si $\lambda_1 = 1$?

16. Verifique que el polinomio característico de la matriz de Leslie L en la ecuación (3) es

$$c_L(\lambda) = (-1)^n(\lambda^n - b_1\lambda^{n-1} - b_2s_1\lambda^{n-2} - b_3s_1s_2\lambda^{n-3} - \cdots - b_ns_1s_2\cdots s_{n-1})$$

[*Sugerencia:* use inducción matemática y desarrolle det$(L - \lambda I)$ a lo largo de la última columna.]

17. Si todas las tasas de supervivencia s_i son distintas de cero, sea

$$P = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & s_1 & 0 & \cdots & 0 \\ 0 & 0 & s_1s_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & s_1s_2\cdots s_{n-1} \end{bmatrix}$$

Calcule $P^{-1}LP$ y úsela para encontrar el polinomio característico de L. [*Sugerencia:* consulte el ejercicio 32 de la sección 4.3.]

18. Verifique que un eigenvector de L correspondiente a λ_1 es

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ s_1/\lambda_1 \\ s_1s_2/\lambda_1^2 \\ s_1s_2s_3/\lambda_1^3 \\ \vdots \\ s_1s_2s_3\cdots s_{n-1}/\lambda_1^{n-1} \end{bmatrix}$$

[*Sugerencia:* combine el ejercicio 17 anterior con el ejercicio 32 de la sección 4.3 y el ejercicio 46 de la sección 4.4.]

CAS *En los ejercicios 19-21, calcule la tasa de crecimiento de estado estacionario de la población con la matriz de Leslie L del ejercicio dado. Luego use el ejercicio 18 para encontrar la correspondiente distribución de las clases etáreas.*

19. Ejercicio 19 de la sección 3.7

20. Ejercicio 20 de la sección 3.7

21. Ejercicio 24 de la sección 3.7

CAS **22.** Muchas especies de focas sufren de caza comercial. Las cazan por su piel, grasa y carne. El comercio de pieles, en particular, redujo algunas poblaciones de focas al punto de la extinción. En la actualidad, las más grandes amenazas para las poblaciones de focas son el declive de los bancos de peces debido a la sobreexplotación, la contaminación, la perturbación del hábitat, los desechos marinos y el exterminio de algunas especies por parte de dueños de pescaderías. Algunas focas se han declarado especies en peligro de extinción; otras especies se administran cuidadosamente. La tabla 4.7 proporciona las tasas de natalidad y supervivencia para el lobo marino del norte, divididas en clases etáreas de 2 años. [Los datos se basan en A. E. York y J. R. Hartley, "Pup Production Following Harvest of Female Northern Fur Seals", *Canadian Journal of Fisheries and Aquatic Science, 38* (1981), pp. 84-90.]

Tabla 4.7

Edad (años)	Tasa de natalidad	Tasa de supervivencia
0–2	0.00	0.91
2–4	0.02	0.88
4–6	0.70	0.85
6–8	1.53	0.80
8–10	1.67	0.74
10–12	1.65	0.67
12–14	1.56	0.59
14–16	1.45	0.49
16–18	1.22	0.38
18–20	0.91	0.27
20–22	0.70	0.17
22–24	0.22	0.15
24–26	0.00	0.00

(a) Construya la matriz de Leslie L para estos datos y calcule el eigenvalor positivo y su correspondiente eigenvector positivo.

(b) A largo plazo, ¿qué porcentaje de lobos marinos estará en cada clase etárea y cuál será la tasa de crecimiento?

El ejercicio 23 muestra que el comportamiento a largo plazo de una población puede determinarse directamente de las entradas de su matriz de Leslie.

23. La *tasa de reproducción neta* de una población se define como

$$r = b_1 + b_2s_1 + b_3s_1s_2 + \cdots + b_ns_1s_2\cdots s_{n-1}$$

donde las b_i son las tasas de natalidad y las s_j son las tasas de supervivencia para la población.

(a) Explique por qué r puede interpretarse como el número promedio de hijas nacidas de una sola hembra durante su vida.

(b) Demuestre que $r = 1$ si y sólo si $\lambda_1 = 1$. (Esto representa *crecimiento poblacional cero.*) [*Sugerencia:* sea

$$g(\lambda) = \frac{b_1}{\lambda} + \frac{b_2s_1}{\lambda^2} + \frac{b_3s_1s_2}{\lambda^3} + \cdots + \frac{b_ns_1s_2\cdots s_{n-1}}{\lambda^n}$$

Demuestre que λ es un eigenvalor de L si y sólo si $g(\lambda) = 1$.]

(c) Suponga que existe un eigenvalor positivo único λ_1, y demuestre que $r < 1$ si y sólo si la población disminuye y $r > 1$ si y sólo si la población aumenta.

Una política de reproducción sustentable es un procedimiento que permite que cierta fracción de una población (representada por un vector de distribución de población **x***) se reproduzca de modo que la población regrese a* **x** *después de un intervalo de tiempo (donde un intervalo de tiempo es la longitud de una clase etárea). Si h es la fracción de cada clase etárea que se reproduce, entonces el procedimiento de reproducción puede expresarse matemáticamente del modo siguiente: si comienza con un vector de población* **x***, después de un intervalo de tiempo se tiene L***x***; la reproducción remueve hL***x***, lo que deja*

$$L\mathbf{x} - hL\mathbf{x} = (1 - h)L\mathbf{x}$$

La sustentabilidad requiere que

$$(1 - h)L\mathbf{x} = \mathbf{x}$$

24. Si λ_1 es el único eigenvalor positivo de una matriz de Leslie L y h es la tasa de reproducción sustentable, pruebe que $h = 1 - 1/\lambda_1$.

CAS **25. (a)** Encuentre la proporción de reproducción sustentable para el caribú de bosque del ejercicio 42 de la sección 3.7.

(b) Con los datos del ejercicio 42 de la sección 3.7, reduzca la manada de caribúes de acuerdo con su respuesta al inciso (a). Verifique que la población regresa a su nivel original después de un intervalo de tiempo.

26. Encuentre la tasa de reproducción sustentable para el lobo marino del ejercicio 22. (Los conservacionistas han tenido que reproducir poblaciones de lobos marinos cuando la sobreexplotación pesquera redujo el suministro de alimento disponible al punto en que los lobos marinos están en peligro de morir de inanición.)

a + bi **27.** Sea L una matriz de Leslie con un eigenvalor positivo único λ_1. Demuestre que si λ es cualquier otro eigenvalor (real o complejo) de L, entonces $|\lambda| \leq \lambda_1$. [*Sugerencia:* escriba $\lambda = r(\cos\theta + i\operatorname{sen}\theta)$ y sustitúyala en la ecuación $g(\lambda) = 1$, como en el inciso (b) del ejercicio 23. Use el Teorema de De Moivre y luego tome la parte real de ambos lados. Le resultará útil la desigualdad del triángulo.]

El Teorema de Perron-Frobenius

En los ejercicios 28-31, encuentre la raíz de Perron y el correspondiente eigenvector de Perron de A.

28. $A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$

29. $A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$

30. $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

31. $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

Es posible demostrar que una matriz de $n \times n$ no negativa es irreductible si y sólo si $(I + A)^{n-1} > O$. En los ejercicios 32-35, use este criterio para determinar si la matriz A es irreduc-

tible. Si A es reducible, encuentre una permutación de sus renglones y columnas que ponga A en la forma por bloques.

$$\left[\begin{array}{c|c} B & C \\ \hline O & D \end{array}\right]$$

32. $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ **33.** $A = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$

34. $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$ **35.** $A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$

36. (a) Si A es la matriz de adyacencia de un grafo G, demuestre que A es irreductible si y sólo si G está conectado. (Un grafo está ***conectado*** si existe una ruta entre cada par de vértices.)

(b) ¿Cuál de los grafos de la sección 4.0 tiene una matriz de adyacencia irreductible? ¿Cuál tiene una matriz de adyacencia primitiva?

37. Sea G un grafo bipartito con matriz de adyacencia A.

(a) Demuestre que A no es primitiva.

(b) Demuestre que si λ es un eigenvalor de A, también lo es $-\lambda$. [*Sugerencia:* use el ejercicio 78 de la sección 3.7 y la partición de un eigenvector para λ de modo que sea compatible con este particionamiento de A. Use esta partición para encontrar un eigenvector para $-\lambda$.]

38. Un grafo se llama ***k-regular*** si k aristas se unen en cada vértice. Sea G un grafo k-regular.

(a) Demuestre que la matriz de adyacencia A de G tiene $\lambda = k$ como eigenvalor. [*Sugerencia:* adapte el Teorema 4.30.]

(b) Demuestre que, si A es primitiva, entonces los otros eigenvalores son menores que k en valor absoluto. (*Sugerencia:* adapte el Teorema 4.31.)

39. Explique los resultados de su exploración en la sección 4.0 a la luz de los ejercicios 36-38 y la sección 4.5.

En el ejercicio 40, el valor absoluto de una matriz $A = [a_{ij}]$ *se define como la matriz* $|A| = [|a_{ij}|]$.

40. Sean A y B matrices de $n \times n$, $\mathbf{x}$ un vector en $\mathbb{R}^n$ y c un escalar. Pruebe las siguientes desigualdades matriciales:

(a) $|cA| = |c||A|$ **(b)** $|A + B| \leq |A| + |B|$

(c) $|A\mathbf{x}| \leq |A||\mathbf{x}|$ **(d)** $|AB| \leq |A||B|$

41. Pruebe que una matriz $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ de 2×2 es reducible si y sólo si $a_{12} = 0$ o $a_{21} = 0$.

42. Sea A una matriz no negativa irreductible tal que $I - A$ sea invertible e $(I - A)^{-1} \geq O$. Sea λ_1 y $\mathbf{v}_1$ la raíz de Perron y el eigenvector de Perron de A.

(a) Pruebe que $0 < \lambda_1 < 1$. [*Sugerencia:* aplique el ejercicio 22 de la sección 4.3 y el Teorema 4.18(b).]

(b) Deduzca a partir de (a) que $\mathbf{v}_1 > A\mathbf{v}_1$.

Relaciones de recurrencia lineal

En los ejercicios 43-46, escriba los primeros seis términos de la sucesión definida por la relación de recurrencia con las condiciones iniciales dadas.

43. $x_0 = 1, x_n = 2x_{n-1}$ para $n \geq 1$

44. $a_1 = 128, a_n = a_{n-1}/2$ para $n \geq 2$

45. $y_0 = 0, y_1 = 1, y_n = y_{n-1} - y_{n-2}$ para $n \geq 2$

46. $b_0 = 1, b_1 = 1, b_n = 2b_{n-1} + b_{n-2}$ para $n \geq 2$

En los ejercicios 47-52, resuelva la relación de recurrencia con las condiciones iniciales dadas.

47. $x_0 = 0, x_1 = 5, x_n = 3x_{n-1} + 4x_{n-2}$ para $n \geq 2$

48. $x_0 = 0, x_1 = 1, x_n = 4x_{n-1} - 3x_{n-2}$ para $n \geq 2$

49. $y_1 = 1, y_2 = 6, y_n = 4y_{n-1} - 4y_{n-2}$ para $n \geq 3$

50. $a_0 = 4, a_1 = 1, a_n = a_{n-1} - a_{n-2}/4$ para $n \geq 2$

51. $b_0 = 0, b_1 = 1, b_n = 2b_{n-1} + 2b_{n-2}$ para $n \geq 2$

a + bi **52.** En la relación de recurrencia en el ejercicio 45. Demuestre que su solución concuerda con la respuesta al ejercicio 45.

53. Complete la demostración del Teorema 4.38(a) al demostrar que si la relación de recurrencia $x_n = ax_{n-1} + bx_{n-2}$ tiene distintos eigenvalores $\lambda_1 \neq \lambda_2$, entonces la solución será de la forma

$$x_n = c_1\lambda_1^n + c_2\lambda_2^n$$

[*Sugerencia:* demuestre que el método del ejemplo 4.40 funciona en general.]

54. Demuestre que para cualquier elección de condiciones iniciales $x_0 = r$ y $x_1 = s$, los escalares c_1 y c_2 pueden encontrarse, como se afirma en el Teorema 4.38(a) y (b).

55. La recurrencia de Fibonacci $f_n = f_{n-1} + f_{n-2}$ tiene la ecuación matricial asociada $\mathbf{x}_n = A\mathbf{x}_{n-1}$, donde

$$\mathbf{x}_n = \begin{bmatrix} f_n \\ f_{n-1} \end{bmatrix} \quad \text{y} \quad A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

(a) Con $f_0 = 0$ y $f_1 = 1$, use inducción matemática para demostrar que

$$A^n = \begin{bmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{bmatrix}$$

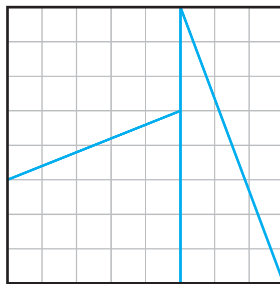
para todo $n \geq 1$.

(b) Con el inciso (a), demuestre que

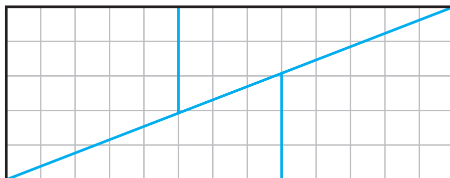
$$f_{n+1}f_{n-1} - f_n^2 = (-1)^n$$

para toda $n \geq 1$. [Esta se llama ***identidad de Cassini***, en honor del astrónomo Giovanni Domenico Cassini (1625-1712). Cassini nació en Italia pero, a invitación de Luis XIV, se mudó a Francia en 1669, donde se convirtió en el director del Observatorio de París. Se convirtió en ciudadano francés y adoptó la versión francesa de su nombre: Jean-Dominique Cassini. Las matemáticas fueron uno de sus principales intereses, además de la astronomía. La Identidad de Cassini se publicó en 1680 en un artículo enviado a la Real Academia de Ciencias de París.]

(c) Un tablero de ajedrez se puede dividir como se muestra en la figura 4.29(a) y las piezas se reacomodan en un rectángulo de 5 × 13 de la figura 4.29(b). El área del cuadrado tiene 64 unidades cuadradas, ¡pero el área del rectángulo es de 65 unidades cuadradas! ¿De dónde provino el cuadrado adicional?



(a)



(b)

Figura 4.29

[*Sugerencia:* ¿qué tiene que ver esto con la sucesión de Fibonacci?]

56. Usted tiene un suministro de tres tipos de mosaicos: dos tipos de mosaicos de 1 × 2 y un tipo de mosaico de 1 × 1, como se muestra en la figura 4.30.



Figura 4.30

Sea t_n el número de diferentes formas de cubrir un rectángulo de $1 \times n$ con estos mosaicos. Por ejemplo, la figura 4.31 muestra que $t_3 = 5$.

(a) Encuentre $t_1, \ldots, t_5$.
(¿Tiene algún sentido t_0? Si es así, ¿cuál es?)

(b) Establezca una relación de recurrencia de segundo orden para t_n.

(c) Con t_1 y t_2 como condiciones iniciales, resuelva la relación de recurrencia en el inciso (b). Verifique su respuesta con los datos del inciso (a).

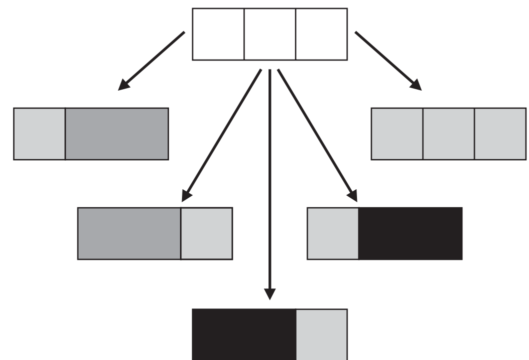


Figura 4.31
Las cinco formas de cubrir un rectángulo de 1 × 3 con los mosaicos

57. Usted tiene un suministro de fichas de dominó de 1 × 2 con las cuales cubrir un rectángulo de $2 \times n$. Sea d_n el número de diferentes formas de cubrir el rectángulo. Por ejemplo, la figura 4.32 muestra que $d_3 = 3$.

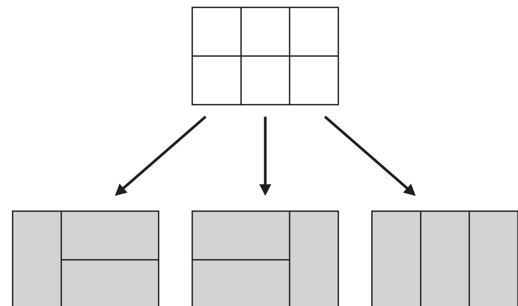


Figura 4.32
Las tres formas de cubrir un rectángulo de 2 × 3 con fichas de dominó de 1 × 2

(a) Encuentre $d_1, \ldots, d_5$.
(¿Tiene algún sentido d_0? Si lo tiene, ¿cuál es?)
(b) Establezca una relación de recurrencia de segundo orden para d_n.
(c) Con d_1 y d_2 como condiciones iniciales, resuelva la relación de recurrencia en el inciso (b). Verifique su respuesta con los datos del inciso (a).

58. En el ejemplo 4.41, encuentre eigenvectores $\mathbf{v}_1$ y $\mathbf{v}_2$ que correspondan a $\lambda_1 = \dfrac{1+\sqrt{5}}{2}$ y $\lambda_2 = \dfrac{1-\sqrt{5}}{2}$. Con $\mathbf{x}_k = \begin{bmatrix} f_k \\ f_{k-1} \end{bmatrix}$, verifique la fórmula (2) de la sección 4.5. Esto es, demuestre que, para algún escalar c_1,

$$\lim_{k\to\infty} \frac{\mathbf{x}_k}{\lambda_1^k} = c_1\mathbf{v}_1$$

Sistemas de ecuaciones diferenciales lineales



En los ejercicios 59-64, encuentre la solución general al sistema dado de ecuaciones diferenciales. Luego encuentre la solución específica que satisfaga las condiciones iniciales. (Considere todas las funciones como funciones de t.)

59. $x' = x + 3y, \quad x(0) = 0$
$y' = 2x + 2y, \quad y(0) = 5$

60. $x' = 2x - y, \quad x(0) = 1$
$y' = -x + 2y, \quad y(0) = 1$

61. $x_1' = x_1 + x_2, \quad x_1(0) = 1$
$x_2' = x_1 - x_2, \quad x_2(0) = 0$

62. $y_1' = y_1 - y_2, \quad y_1(0) = 1$
$y_2' = y_1 + y_2, \quad y_2(0) = 1$

63. $x' = y - z, \quad x(0) = 1$
$y' = x + z, \quad y(0) = 0$
$z' = x + y, \quad z(0) = -1$

64. $x' = x + 3z, \quad x(0) = 2$
$y' = x - 2y + z, \quad y(0) = 3$
$z' = 3x + z, \quad z(0) = 4$

65. Un científico coloca dos cepas de bacterias, X y Y, en una caja de Petri. Inicialmente, existen 400 de X y 500 de Y. Las dos bacterias compiten por alimento y espacio pero no se alimentan una de la otra. Si $x = x(t)$ y $y = y(t)$ son los números de las cepas en el tiempo t días, las tasas de crecimiento de las dos poblaciones están dadas por el sistema

$$\begin{aligned} x' &= 1.2x - 0.2y \\ y' &= -0.2x + 1.5y \end{aligned}$$

(a) Determine qué le ocurre a estas dos poblaciones al resolver el sistema de ecuaciones diferenciales.
(b) Explore el efecto de cambiar las poblaciones iniciales al hacer $x(0) = a$ y $y(0) = b$. Describa lo que ocurre con las poblaciones en términos de a y b.

66. Dos especies, X y Y, viven en una relación *simbiótica*. Esto es, ninguna especie puede sobrevivir por cuenta propia y cada una depende de la otra para su supervivencia. Inicialmente, hay 15 de X y 10 de Y. Si $x = x(t)$ y $y = y(t)$ son los tamaños de las poblaciones en el tiempo t meses, las tasas de crecimiento de las dos poblaciones están dadas por el sistema

$$\begin{aligned} x' &= -0.8x + 0.4y \\ y' &= 0.4x - 0.2y \end{aligned}$$

Determine lo que le pasa a estas dos poblaciones.

En los ejercicios 67 y 68, la especie X depreda a la especie Y. Los tamaños de las poblaciones se representan mediante $x = x(t)$ y $y = y(t)$. La tasa de crecimiento de cada población está gobernada por el sistema de ecuaciones diferenciales $\mathbf{x}' = A\mathbf{x} + \mathbf{b}$, donde $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ y $\mathbf{b}$ es un vector constante. Determine qué le ocurre a las dos poblaciones para los A y $\mathbf{b}$ dados y las condiciones iniciales $\mathbf{x}(0)$. (Primero demuestre que existen constantes a y b tales que las sustituciones $x = u + a$ y $y = v + b$ convierten el sistema en uno equivalente sin términos constantes.)

67. $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} -30 \\ -10 \end{bmatrix}, \mathbf{x}(0) = \begin{bmatrix} 20 \\ 30 \end{bmatrix}$

68. $A = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 0 \\ 40 \end{bmatrix}, \mathbf{x}(0) = \begin{bmatrix} 10 \\ 30 \end{bmatrix}$

69. Sea $x = x(t)$ una función dos veces derivable y considere la *ecuación diferencial de segundo orden*

$$x'' + ax' + bx = 0 \tag{11}$$

(a) Demuestre que el cambio de variables $y = x'$ y $z = x$ permite a la ecuación (11) escribirse como un sistema de dos ecuaciones diferenciales lineales en y y z.
(b) Demuestre que la ecuación característica del sistema en el inciso (a) es $\lambda^2 + a\lambda + b = 0$.

70. Demuestre que hay un cambio de variables que convierte la *ecuación diferencial de n-ésimo orden*

$$x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_1x' + a_0 = 0$$

en un sistema de n ecuaciones diferenciales lineales cuya matriz de coeficientes es la matriz acompañante $C(p)$ del polinomio $p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0$. [La notación $x^{(k)}$ denota la k-ésima derivada de x. Vea en los ejercicios 26-32 de la sección 4.3 la definición de una matriz acompañante.]

En los ejercicios 71 y 72, use el ejercicio 69 para encontrar la solución general de la ecuación dada.

71. $x'' - 5x' + 6x = 0$ **72.** $x'' + 4x' + 3x = 0$

En los ejercicios 73-76, resuelva el sistema de ecuaciones diferenciales en el ejercicio dado usando el Teorema 4.41.

73. Ejercicio 59 **74.** Ejercicio 60

75. Ejercicio 63 **76.** Ejercicio 64

Sistemas dinámicos lineales discretos

En los ejercicios 77-84, considere el sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

(a) *Calcule y grafique* $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ para $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(b) *Calcule y grafique* $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ para $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

(c) *Con eigenvalores y eigenvectores, clasifique el origen como atractor, repelente, punto de silla o ninguno de los anteriores.*

(d) *Bosqueje varias trayectorias típicas del sistema.*

77. $A = \begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$ **78.** $A = \begin{bmatrix} 0.5 & -0.5 \\ 0 & 0.5 \end{bmatrix}$

79. $A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ **80.** $A = \begin{bmatrix} -4 & 2 \\ 1 & -3 \end{bmatrix}$

81. $A = \begin{bmatrix} 1.5 & -1 \\ -1 & 0 \end{bmatrix}$ **82.** $A = \begin{bmatrix} 0.1 & 0.9 \\ 0.5 & 0.5 \end{bmatrix}$

83. $A = \begin{bmatrix} 0.2 & 0.4 \\ -0.2 & 0.8 \end{bmatrix}$ **84.** $A = \begin{bmatrix} 0 & -1.5 \\ 1.2 & 3.6 \end{bmatrix}$

En los ejercicios 85-88, la matriz dada es de la forma $A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. *En cada caso, A puede factorizarse como el producto de una matriz de escalamiento y una matriz de rotación. Encuentre el factor de escalamiento r y el ángulo de rotación θ. Grafique los primeros cuatro puntos de la trayectoria para el sistema dinámico* $\mathbf{x}_{k+1} = A\mathbf{x}_k$ *con* $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *y clasifique el origen como un atractor espiral, un repelente espiral o un centro orbital.*

85. $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ **86.** $A = \begin{bmatrix} 0 & 0.5 \\ -0.5 & 0 \end{bmatrix}$

87. $A = \begin{bmatrix} 1 & \sqrt{3} \\ -\sqrt{3} & 1 \end{bmatrix}$ **88.** $A = \begin{bmatrix} -\sqrt{3}/2 & -1/2 \\ 1/2 & -\sqrt{3}/2 \end{bmatrix}$

En los ejercicios 89-92, encuentre una matriz invertible P y una matriz C de la forma $C = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ *tal que* $A = PCP^{-1}$. *Grafique los primeros seis puntos de la trayectoria para el sistema dinámico* $\mathbf{x}_{k+1} = A\mathbf{x}_k$ *con* $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ *y clasifique el origen como atractor espiral, repelente espiral o centro orbital.*

89. $A = \begin{bmatrix} 0.1 & -0.2 \\ 0.1 & 0.3 \end{bmatrix}$ **90.** $A = \begin{bmatrix} 2 & 1 \\ -2 & 0 \end{bmatrix}$

91. $A = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$ **92.** $A = \begin{bmatrix} 0 & -1 \\ 1 & \sqrt{3} \end{bmatrix}$

Repaso del capítulo

Definiciones y conceptos clave

Preguntas de repaso

1. Marque cada uno de los siguientes enunciados como verdadero o falso:

(a) Para toda matriz cuadrada A, $\det(-A) = -\det A$.

(b) Si A y B son matrices de $n \times n$, entonces $\det(AB) = \det(BA)$.
(c) Si A y B son matrices de $n \times n$ cuyas columnas son iguales pero en diferente orden, entonces $\det B = -\det A$.
(d) Si A es invertible, entonces $\det(A^{-1}) = \det A^T$.
(e) Si 0 es el único eigenvalor de una matriz cuadrada A, entonces A es la matriz cero.
(f) Dos eigenvectores que corresponden al mismo eigenvalor deben ser linealmente dependientes.
(g) Si una matriz de $n \times n$ tiene n eigenvalores distintos, entonces debe ser diagonalizable.
(h) Si una matriz de $n \times n$ es diagonalizable, entonces debe tener n eigenvalores distintos.
(i) Las matrices semejantes tienen los mismos eigenvectores.
(j) Si A y B son dos matrices de $n \times n$ con la misma forma escalonada reducida por renglones, entonces A es semejante a B.

2. Sea $A = \begin{bmatrix} 1 & 3 & 5 \\ 3 & 5 & 7 \\ 7 & 9 & 11 \end{bmatrix}$.

(a) Calcule $\det A$ mediante expansión por cofactores a lo largo de cualquier renglón o columna.
(b) Calcule $\det A$ al reducir primero A a la forma triangular.

3. Si $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = 3$, encuentre $\begin{vmatrix} 3d & 2e - 4f & f \\ 3a & 2b - 4c & c \\ 3g & 2h - 4i & i \end{vmatrix}$.

4. Sean A y B matrices de 4×4 con $\det A = 2$ y $\det B = -\frac{1}{4}$. Encuentre $\det C$ para la matriz C indicada:
(a) $C = (AB)^{-1}$ **(b)** $C = A^2B(3A^T)$

5. Si A es una matriz de $n \times n$ antisimétrica y n es impar, pruebe que $\det A = 0$.

6. Encuentre todos los valores de k para los cuales

$$\begin{vmatrix} 1 & -1 & 2 \\ 1 & 1 & k \\ 2 & 4 & k^2 \end{vmatrix} = 0.$$

En las preguntas 7 y 8, demuestre que **x** *es un eigenvector de A y encuentre el correspondiente eigenvalor.*

7. $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, A = \begin{bmatrix} 3 & 1 \\ 4 & 3 \end{bmatrix}$

8. $\mathbf{x} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}, A = \begin{bmatrix} 13 & -60 & -45 \\ -5 & 18 & 15 \\ 10 & -40 & -32 \end{bmatrix}$

9. Sea $A = \begin{bmatrix} -5 & -6 & 3 \\ 3 & 4 & -3 \\ 0 & 0 & -2 \end{bmatrix}$.

(a) Encuentre el polinomio característico de A.
(b) Encuentre todos los eigenvalores de A.
(c) Encuentre una base para cada uno de los eigenespacios de A.
(d) Determine si A es diagonalizable. Si A no es diagonalizable, explique por qué no. Si A es diagonalizable, encuentre una matriz invertible P y una matriz diagonal D tales que $P^{-1}AP = D$.

10. Si A es una matriz diagonalizable de 3×3 con eigenvalores -2, 3 y 4, encuentre $\det A$.

11. Si A es una matriz de 2×2 con eigenvalores $\lambda_1 = \frac{1}{2}$, $\lambda_2 = -1$, y correspondientes eigenvectores

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \text{ encuentre } A^{-5}\begin{bmatrix} 3 \\ 7 \end{bmatrix}.$$

12. Si A es una matriz diagonalizable y todos sus eigenvalores satisfacen $|\lambda| < 1$, demuestre que A^n tiende a la matriz cero conforme n se vuelve grande.

En las preguntas 13-15, determine, con razones, si A es semejante a B. Si A ~ B, proporcione una matriz invertible P tal que $P^{-1}AP = B$.

13. $A = \begin{bmatrix} 4 & 2 \\ 3 & 1 \end{bmatrix}, B = \begin{bmatrix} 2 & 2 \\ 3 & 2 \end{bmatrix}$

14. $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}, B = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$

15. $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

16. Sea $A = \begin{bmatrix} 2 & k \\ 0 & 1 \end{bmatrix}$. Encuentre todos los valores de k para los cuales:
(a) A tiene eigenvalores 3 y -1.
(b) A tiene un eigenvalor con multiplicidad algebraica 2.
(c) A no tiene eigenvalores reales.

17. Si $A^3 = A$, ¿cuáles son los posibles eigenvalores de A?

18. Si una matriz cuadrada A tiene dos renglones iguales, ¿por qué A debe tener al 0 como uno de sus eigenvalores?

19. Si **x** es un eigenvector de A con eigenvalor $\lambda = 3$, demuestre que **x** también es un eigenvector de $A^2 - 5A + 2I$. ¿Cuál es el eigenvalor correspondiente?

20. Si A es semejante a B, con $P^{-1}AP = B$, y **x** es un eigenvector de A, demuestre que $P^{-1}\mathbf{x}$ es un eigenvector de B.

5 Ortogonalidad

. . . ese vivaz escocés de todos los escoceses, Douglas, que trepa a caballo una colina perpendicular. . .

—William Shakespeare
Enrique IV, Parte I, Acto II, Escena IV

5.0 Introducción: sombras en la pared

En este capítulo se extenderá la noción de proyección ortogonal que encontró por primera vez en el capítulo 1 y luego nuevamente en el capítulo 3. Hasta el momento sólo se han estudiado proyecciones sobre un solo vector (o, de manera equivalente, el subespacio unidimensional generado por dicho vector). En esta sección se verá si es posible encontrar las fórmulas análogas para la proyección sobre un plano en $\mathbb{R}^3$. La figura 5.1 muestra lo que ocurre, por ejemplo, cuando rayos de luz paralelos crean una sombra en una pared. Un proceso similar ocurre cuando un objeto tridimensional se despliega sobre una pantalla bidimensional, como un monitor de computadora. Más adelante en este capítulo se considerarán estas ideas en toda su generalidad.

Para comenzar, recuerde lo que ya sabe acerca de las proyecciones. En la sección 3.6 se demostró que, en $\mathbb{R}^2$, la matriz estándar de una proyección sobre la recta que pasa por el origen con vector director $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$ es

$$P = \frac{1}{d_1^2 + d_2^2}\begin{bmatrix} d_1^2 & d_1d_2 \\ d_1d_2 & d_2^2 \end{bmatrix} = \begin{bmatrix} d_1^2/(d_1^2 + d_2^2) & d_1d_2/(d_1^2 + d_2^2) \\ d_1d_2/(d_1^2 + d_2^2) & d_2^2/(d_1^2 + d_2^2) \end{bmatrix}$$

Por tanto, la proyección del vector $\mathbf{v}$ sobre esta recta es justo $P\mathbf{v}$.

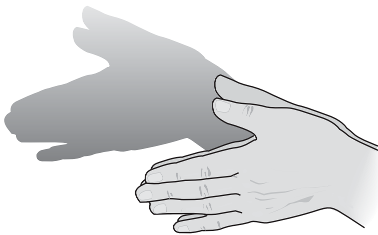


Figura 5.1
Las sombras sobre una pared son proyecciones

Problema 1 Demuestre que P puede escribir en la forma equivalente

$$P = \begin{bmatrix} \cos^2\theta & \cos\theta\,\text{sen}\,\theta \\ \cos\theta\,\text{sen}\,\theta & \text{sen}^2\theta \end{bmatrix}$$



(¿Qué representa θ aquí?)

Problema 2 Demuestre que P también puede escribirse en la forma $P = \mathbf{u}\mathbf{u}^T$, donde $\mathbf{u}$ es un vector *unitario* en la dirección de $\mathbf{d}$.

Problema 3 Con el problema 2, encuentre P y luego encuentre la proyección de $\mathbf{v} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}$ sobre las rectas con los siguientes vectores directores unitarios:

$$\text{(a)}\ \mathbf{u} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \quad \text{(b)}\ \mathbf{u} = \begin{bmatrix} \frac{4}{5} \\ \frac{3}{5} \end{bmatrix} \quad \text{(c)}\ \mathbf{u} = \begin{bmatrix} -\frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$$

Problema 4 Con la forma $P = \mathbf{u}\mathbf{u}^T$, demuestre que (a) $P^T = P$ (esto es, P es simétrica) y (b) $P^2 = P$ (esto es, P es idempotente).

Problema 5 Explique por qué si P es una matriz de proyección de 2×2, la recta sobre la que proyecta vectores es el espacio columna de P.

Ahora avance hacia $\mathbb{R}^3$ y considere las proyecciones sobre los planos a través del origen. Se explorarán varios enfoques.

La figura 5.2 muestra una forma de proceder. Si $\mathscr{P}$ es un plano que pasa por el origen en $\mathbb{R}^3$ con vector normal $\mathbf{n}$, y si $\mathbf{v}$ es un vector en $\mathbb{R}^3$, entonces $\mathbf{p} = \text{proy}_{\mathscr{P}}(\mathbf{v})$ es un vector en $\mathscr{P}$ tal que $\mathbf{v} - c\mathbf{n} = \mathbf{p}$ para algún escalar c.

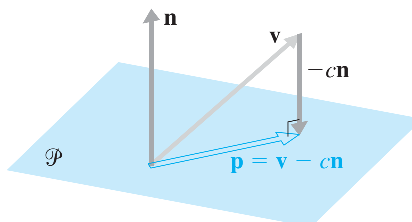


Figura 5.2
Proyección sobre un plano

Problema 6 Con el hecho de que $\mathbf{n}$ es ortogonal a todo vector en $\mathscr{P}$, resuelva $\mathbf{v} - c\mathbf{n} = \mathbf{p}$ para c y encuentre una expresión para $\mathbf{p}$ en términos de $\mathbf{v}$ y $\mathbf{n}$.

Problema 7 Use el método del problema 6 para encontrar la proyección de

$$\mathbf{v} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}$$

sobre los planos con las siguientes ecuaciones:

$$\text{(a)}\ x + y + z = 0 \quad \text{(b)}\ x - 2z = 0 \quad \text{(c)}\ 2x - 3y + z = 0$$

Otro enfoque del problema de encontrar la proyección de un vector sobre un plano se sugiere a partir de la figura 5.3. Puede descomponer la proyección de $\mathbf{v}$ sobre $\mathscr{P}$ en la *suma* de sus proyecciones sobre los vectores directores para $\mathscr{P}$. Esto funciona sólo si los vectores directores son unitarios ortogonales. En concordancia, sean $\mathbf{u}_1$ y $\mathbf{u}_2$ vectores directores para $\mathscr{P}$ con la propiedad de que

$$\|\mathbf{u}_1\| = \|\mathbf{u}_2\| = 1 \quad \text{y} \quad \mathbf{u}_1 \cdot \mathbf{u}_2 = 0$$

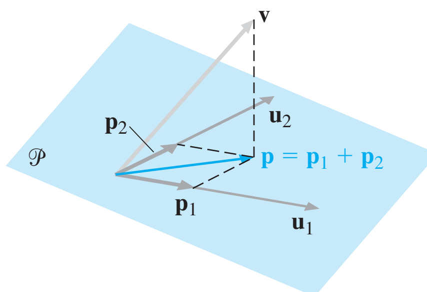


Figura 5.3

Por el problema 2, las proyecciones de $\mathbf{v}$ sobre $\mathbf{u}_1$ y $\mathbf{u}_2$ son

$$\mathbf{p}_1 = \mathbf{u}_1\mathbf{u}_1^T\mathbf{v} \quad \text{y} \quad \mathbf{p}_2 = \mathbf{u}_2\mathbf{u}_2^T\mathbf{v}$$

respectivamente. Para demostrar que $\mathbf{p}_1 + \mathbf{p}_2$ produce la proyección de $\mathbf{v}$ sobre $\mathcal{P}$, es necesario demostrar que $\mathbf{v} - (\mathbf{p}_1 + \mathbf{p}_2)$ es ortogonal al plano $\mathcal{P}$. Es suficiente demostrar que $\mathbf{v} - (\mathbf{p}_1 + \mathbf{p}_2)$ es ortogonal tanto a $\mathbf{u}_1$ como a $\mathbf{u}_2$. (¿Por qué?)

Problema 8 Demuestre que $\mathbf{u}_1 \cdot (\mathbf{v} - (\mathbf{p}_1 + \mathbf{p}_2)) = 0$ y $\mathbf{u}_2 \cdot (\mathbf{v} - (\mathbf{p}_1 + \mathbf{p}_2)) = 0$. [*Sugerencia:* use la forma alternativa del producto punto, $\mathbf{x}^T\mathbf{y} = \mathbf{x} \cdot \mathbf{y}$, junto con el hecho de que $\mathbf{u}_1$ y $\mathbf{u}_2$ son vectores unitarios ortogonales.]

Del problema 8, y de los comentarios que lo anteceden se tiene que la matriz de la proyección sobre el subespacio $\mathcal{P}$ de $\mathbb{R}^3$ generado por los vectores unitarios ortogonales $\mathbf{u}_1$ y $\mathbf{u}_2$ es

$$P = \mathbf{u}_1\mathbf{u}_1^T + \mathbf{u}_2\mathbf{u}_2^T \tag{1}$$

Problema 9 Repita el problema 7 usando la fórmula para P dada por la ecuación (1). Use el mismo $\mathbf{v}$ y use $\mathbf{u}_1$ y $\mathbf{u}_2$, como se indica a continuación. (Primero, verifique que $\mathbf{u}_1$ y $\mathbf{u}_2$ son vectores unitarios ortogonales en el plano dado.)

(a) $x + y + z = 0$ con $\mathbf{u}_1 = \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$ y $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$

(b) $x - 2z = 0$ con $\mathbf{u}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 0 \\ 1/\sqrt{5} \end{bmatrix}$ y $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

(c) $2x - 3y + z = 0$ con $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$ y $\mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$

Problema 10 Demuestre que una matriz de proyección dada por la ecuación (1) satisface las propiedades (a) y (b) del problema 4.

Problema 11 Demuestre que la matriz P de una proyección sobre un plano en $\mathbb{R}^3$ puede expresarse como

$$P = AA^T$$

para alguna matriz A de 3×2. [*Sugerencia:* demuestre que la ecuación (1) es una expansión de producto exterior.]

Problema 12 Demuestre que si P es la matriz de una proyección sobre un plano en $\mathbb{R}^3$, entonces rango$(P) = 2$.

En este capítulo se analizarán los conceptos de ortogonalidad y proyección ortogonal con mayor detalle. Se verá que las ideas introducidas en esta sección pueden generalizarse y que tienen muchas aplicaciones importantes.

5.1 Ortogonalidad en $\mathbb{R}^n$

En esta sección se generalizará la noción de ortogonalidad de vectores en $\mathbb{R}^n$ desde dos vectores hasta conjuntos de vectores. Para hacerlo, se verá que dos propiedades facilitarán el trabajo con las bases estándar $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de $\mathbb{R}^n$: primera, dos vectores distintos cualesquiera en el conjunto son ortogonales. Segunda, cada vector en el conjunto es

unitario. Estas dos propiedades conducen a la noción de bases ortogonales y bases ortonormales, conceptos que podrán aplicarse fructíferamente a varias aplicaciones.

Conjuntos ortogonales y ortonormales de vectores

Definición Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ en $\mathbb{R}^n$ se llama ***conjunto ortogonal*** si todos los pares de vectores distintos en el conjunto son ortogonales; esto es, si

$$\mathbf{v}_i \cdot \mathbf{v}_j = 0 \quad \text{siempre que} \quad i \neq j \quad \text{para} \quad i, j = 1, 2, \ldots, k$$

La base estándar $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ de $\mathbb{R}^n$ es un conjunto ortogonal, como lo es cualquier subconjunto del mismo. Como ilustra el primer ejemplo, existen muchas otras posibilidades.

Ejemplo 5.1

Demuestre que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es un conjunto ortogonal en $\mathbb{R}^3$ si

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Solución Debe demostrar que todo par de vectores de este conjunto es ortogonal. Esto es verdadero, pues

$$\begin{aligned} \mathbf{v}_1 \cdot \mathbf{v}_2 &= 2(0) + 1(1) + (-1)(1) = 0 \\ \mathbf{v}_2 \cdot \mathbf{v}_3 &= 0(1) + 1(-1) + (1)(1) = 0 \\ \mathbf{v}_1 \cdot \mathbf{v}_3 &= 2(1) + 1(-1) + (-1)(1) = 0 \end{aligned}$$

Geométricamente, los vectores del ejemplo 5.1 son mutuamente perpendiculares, como lo muestra la figura 5.4.

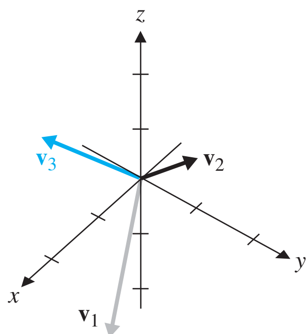


Figura 5.4
Un conjunto ortogonal de vectores

Una de las principales ventajas de trabajar con conjuntos ortogonales de vectores es que necesariamente son linealmente independientes, como lo muestra el Teorema 5.1.

Teorema 5.1 Si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un conjunto ortogonal de vectores distintos de cero en $\mathbb{R}^n$, entonces dichos vectores son linealmente independientes.

Demostración Si $c_1, \ldots, c_k$ son escalares tales que $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$, entonces

$$(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i = \mathbf{0} \cdot \mathbf{v}_i = 0$$

o, de manera equivalente,

$$c_1(\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + c_i(\mathbf{v}_i \cdot \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k \cdot \mathbf{v}_i) = 0 \qquad (2)$$

Dado que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un conjunto ortogonal, todos los productos punto en la ecuación (2) son cero, excepto $\mathbf{v}_i \cdot \mathbf{v}_i$. Por tanto, la ecuación (2) se reduce a

$$c_i(\mathbf{v}_i \cdot \mathbf{v}_i) = 0$$

Ahora, $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$ porque $\mathbf{v}_i \neq \mathbf{0}$, por hipótesis. Así que debe tener $c_i = 0$. El hecho de que esto sea cierto para todo $i = 1, \ldots, k$ implica que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un conjunto linealmente independiente.

Comentario Gracias al Teorema 5.1 se sabe que, si un conjunto de vectores es ortogonal, automáticamente es linealmente independiente. Por ejemplo, inmediatamente puede deducir que los tres vectores en el ejemplo 5.1 son linealmente independientes. ¡Compare este planteamiento con el trabajo necesario para establecer directamente su independencia lineal!

Definición Una ***base ortogonal*** para un subespacio W de $\mathbb{R}^n$ es una base de W que es un conjunto ortogonal.

Ejemplo 5.2

Los vectores

$$\mathbf{v}_1 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

del ejemplo 5.1 son ortogonales y, por ende, linealmente independientes. Dado que cualesquiera tres vectores linealmente independientes en $\mathbb{R}^3$ forman una base para $\mathbb{R}^3$, por el teorema fundamental de las matrices invertibles, se tiene que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortogonal para $\mathbb{R}^3$.

Comentario En el ejemplo 5.2, suponga que sólo se proporcionaron los vectores ortogonales $\mathbf{v}_1$ y $\mathbf{v}_2$, y que se le pidió encontrar un tercer vector $\mathbf{v}_3$ para hacer a $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ una base ortogonal para $\mathbb{R}^3$. Una forma de hacer esto es recordar que en $\mathbb{R}^3$ el producto cruz de dos vectores $\mathbf{v}_1$ y $\mathbf{v}_2$ es ortogonal a cada uno de ellos. (Vea la Exploración: "El producto cruz", en el capítulo 1.) Por tanto, se tiene

$$\mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} \times \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ 2 \end{bmatrix}$$

Note que el vector resultante es un múltiplo del vector $\mathbf{v}_3$ en el ejemplo 5.2, como debe ser.

Ejemplo 5.3

Encuentre una base ortogonal para el subespacio W de $\mathbb{R}^3$ dado por

$$W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x - y + 2z = 0 \right\}$$

Solución La sección 5.3 ofrece un procedimiento general para los problemas de este tipo. Por ahora, se encontrará la base ortogonal mediante fuerza bruta. El subespacio W es un plano que pasa por el origen en $\mathbb{R}^3$. A partir de la ecuación del plano se tiene $x = y - 2z$, de modo que W consiste de vectores de la forma

$$\begin{bmatrix} y - 2z \\ y \\ z \end{bmatrix} = y\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

Se tiene que $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ son una base para W, pero *no* son ortogonales. Es suficiente encontrar otro vector distinto de cero en W que sea ortogonal a cualquiera de ellos.

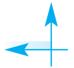
Suponga que $\mathbf{w} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ es un vector en W que es ortogonal a $\mathbf{u}$. Entonces $x - y + 2z = 0$, pues $\mathbf{w}$ está en el plano W. Dado que $\mathbf{u} \cdot \mathbf{w} = 0$, también se tiene $x + y = 0$. Al resolver el sistema lineal

$$\begin{aligned} x - y + 2z &= 0 \\ x + y &= 0 \end{aligned}$$

se encuentra que $x = -z$ y $y = z$. (Compruebe esto.) Por tanto, cualquier vector $\mathbf{w}$ distinto de cero de la forma

$$\mathbf{w} = \begin{bmatrix} -z \\ z \\ z \end{bmatrix}$$

será adecuado. Para ser específicos, podría tomar $\mathbf{w} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$. Es fácil comprobar que $\{\mathbf{u}, \mathbf{w}\}$ es un conjunto ortogonal en W y, por tanto, una base ortogonal para W, pues $\dim W = 2$.



Otra ventaja de trabajar con una base ortogonal es que las coordenadas de un vector con respecto a tal base son fáciles de calcular. De hecho, hay una fórmula para dichas coordenadas, como establece el siguiente teorema.

Teorema 5.2 Sea $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ una base ortogonal para un subespacio W de $\mathbb{R}^n$ y sea $\mathbf{w}$ cualquier vector en W. Entonces los escalares únicos $c_1, \ldots, c_k$ tales que

$$\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

están dados por

$$c_i = \frac{\mathbf{w} \cdot \mathbf{v}_i}{\mathbf{v}_i \cdot \mathbf{v}_i} \quad \text{para } i = 1, \ldots, k$$

Demostración Dado que $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es una base para W, se sabe que existen escalares únicos $c_1, \ldots, c_k$ tales que $\mathbf{w} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ (del Teorema 3.29). Para establecer la fórmula para c_i, se toma el producto punto de esta combinación lineal con $\mathbf{v}_i$ para obtener

$$\begin{aligned} \mathbf{w} \cdot \mathbf{v}_i &= (c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) \cdot \mathbf{v}_i \\ &= c_1(\mathbf{v}_1 \cdot \mathbf{v}_i) + \cdots + c_i(\mathbf{v}_i \cdot \mathbf{v}_i) + \cdots + c_k(\mathbf{v}_k \cdot \mathbf{v}_i) \\ &= c_i(\mathbf{v}_i \cdot \mathbf{v}_i) \end{aligned}$$

pues $\mathbf{v}_j \cdot \mathbf{v}_i = 0$ para $j \neq i$. Dado que $\mathbf{v}_i \neq \mathbf{0}$, $\mathbf{v}_i \cdot \mathbf{v}_i \neq 0$. Al dividir entre $\mathbf{v}_i \cdot \mathbf{v}_i$, se obtiene el resultado deseado.

Ejemplo 5.4

Encuentre las coordenadas de $\mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ con respecto a la base ortogonal $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ de los ejemplos 5.1 y 5.2.

Solución Con el Teorema 5.2, calcule

$$c_1 = \frac{\mathbf{w} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} = \frac{2 + 2 - 3}{4 + 1 + 1} = \frac{1}{6}$$

$$c_2 = \frac{\mathbf{w} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} = \frac{0 + 2 + 3}{0 + 1 + 1} = \frac{5}{2}$$

$$c_3 = \frac{\mathbf{w} \cdot \mathbf{v}_3}{\mathbf{v}_3 \cdot \mathbf{v}_3} = \frac{1 - 2 + 3}{1 + 1 + 1} = \frac{2}{3}$$

Por tanto,

$$\mathbf{w} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 = \tfrac{1}{6}\mathbf{v}_1 + \tfrac{5}{2}\mathbf{v}_2 + \tfrac{2}{3}\mathbf{v}_3$$



(Compruebe esto.) Con la notación introducida en la sección 3.5, también puede escribir la ecuación anterior como

$$[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} \frac{1}{6} \\ \frac{5}{2} \\ \frac{2}{3} \end{bmatrix}$$

Compare el procedimiento del ejemplo 5.4 con el trabajo requerido para encontrar directamente estas coordenadas y comience a apreciar el valor de las bases ortogonales.

Como se anotó al comienzo de esta sección, la otra propiedad de la base estándar en $\mathbb{R}^n$ es que cada vector de base estándar es un vector unitario. Al combinar esta propiedad con la ortogonalidad, se tiene la siguiente definición.

Definición Un conjunto de vectores en $\mathbb{R}^n$ es un ***conjunto ortonormal*** si es un conjunto ortogonal de vectores unitarios. Una ***base ortonormal*** para un subespacio W de $\mathbb{R}^n$ es una base de W que es un conjunto ortonormal.

Comentario Si $S = \{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$ es un conjunto ortonormal de vectores, entonces $\mathbf{q}_i \cdot \mathbf{q}_j = 0$ para $i \neq j$ y $\|\mathbf{q}_i\| = 1$. El hecho de que cada $\mathbf{q}_i$ es un vector unitario es equivalente a $\mathbf{q}_i \cdot \mathbf{q}_i = 1$. Se tiene que es posible resumir el enunciado de que S es ortonormal como

$$\mathbf{q}_i \cdot \mathbf{q}_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

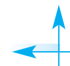
Ejemplo 5.5

Demuestre que $S = \{\mathbf{q}_1, \mathbf{q}_2\}$ es un conjunto ortonormal en $\mathbb{R}^3$ si

$$\mathbf{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \quad \text{y} \quad \mathbf{q}_2 = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

Solución Compruebe que

$$\mathbf{q}_1 \cdot \mathbf{q}_2 = 1/\sqrt{18} - 2/\sqrt{18} + 1/\sqrt{18} = 0$$
$$\mathbf{q}_1 \cdot \mathbf{q}_1 = 1/3 + 1/3 + 1/3 = 1$$
$$\mathbf{q}_2 \cdot \mathbf{q}_2 = 1/6 + 4/6 + 1/6 = 1$$



Si tiene un conjunto ortogonal, puede obtener fácilmente un conjunto ortonormal a partir de él: simplemente normalice cada vector.

Ejemplo 5.6 Construya una base ortonormal para $\mathbb{R}^3$ a partir de los vectores del ejemplo 5.1.

Solución Dado que ya se sabe que $\mathbf{v}_1$, $\mathbf{v}_2$ y $\mathbf{v}_3$ son una base ortogonal, normalícelos para obtener

$$\mathbf{q}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{6}}\begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$$

$$\mathbf{q}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\mathbf{q}_3 = \frac{1}{\|\mathbf{v}_3\|}\mathbf{v}_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Entonces $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ es una base ortonormal para $\mathbb{R}^3$.



Dado que cualquier conjunto ortonormal de vectores es, en particular, ortogonal, entonces es linealmente independiente, por el Teorema 5.1. Si tiene una base ortonormal, el Teorema 5.2 se vuelve todavía más simple.

Teorema 5.3 Sea $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_k\}$ una base ortonormal para un subespacio W de $\mathbb{R}^n$ y sea $\mathbf{w}$ cualquier vector en W. Entonces

$$\mathbf{w} = (\mathbf{w} \cdot \mathbf{q}_1)\mathbf{q}_1 + (\mathbf{w} \cdot \mathbf{q}_2)\mathbf{q}_2 + \cdots + (\mathbf{w} \cdot \mathbf{q}_k)\mathbf{q}_k$$

y esta representación es única.

Demostración Aplique el Teorema 5.2 y use el hecho de que $\mathbf{q}_i \cdot \mathbf{q}_i = 1$ para $i = 1, \ldots, k$.

Matrices ortogonales

Las matrices cuyas columnas forman un conjunto ortonormal surgen frecuentemente en las aplicaciones, como verá en la sección 5.5. Tales matrices tienen muchas propiedades atractivas, que se examinan a continuación.

Teorema 5.4 Las columnas de una matriz Q de $m \times n$ forman un conjunto ortonormal si y sólo si $Q^TQ = I_n$.

Demostración Es necesario demostrar que

$$(Q^TQ)_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Sea $\mathbf{q}_i$ que denota la i-ésima columna de Q (y, en consecuencia, el i-ésimo renglón de Q^T). Dado que la entrada (i, j) de Q^TQ es el producto punto del i-ésimo renglón de Q^T y la j-ésima columna de Q, se tiene que

$$(Q^TQ)_{ij} = \mathbf{q}_i \cdot \mathbf{q}_j \tag{2}$$

por la definición de multiplicación de matrices.

Ahora las columnas Q forman un conjunto ortonormal si y sólo si

$$\mathbf{q}_i \cdot \mathbf{q}_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

lo que, por la ecuación (2), se cumple si y sólo si

$$(Q^TQ)_{ij} = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Esto completa la demostración.

Matriz ortogonal es una desafortunada expresión terminológica. "Matriz ortonormal" sería claramente un mejor término, pero no es de uso común. Más aún, no existe un término para una matriz no cuadrada con columnas ortonormales.

Si la matriz Q en el Teorema 5.4 es *cuadrada*, tiene un nombre especial.

Definición Una matriz Q de $n \times n$ cuyas columnas forman un conjunto ortonormal se llama ***matriz ortogonal***.

El hecho más importante acerca de las matrices ortogonales está dado por el siguiente teorema.

Teorema 5.5 Una matriz cuadrada Q es ortogonal si y sólo si $Q^{-1} = Q^T$.

Demostración Por el Teorema 5.4, Q es ortogonal si y sólo si $Q^TQ = I$. Esto es verdadero si y sólo Q es invertible y $Q^{-1} = Q^T$, por el Teorema 3.13.

Ejemplo 5.7 Demuestre que las siguientes matrices son ortogonales y encuentre sus inversas:

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

Solución Las columnas de A son justo los vectores base estándar para $\mathbb{R}^3$, que claramente son ortonormales. Por tanto, A es ortogonal y

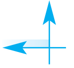
$$A^{-1} = A^T = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Para B, se comprueba directamente que

$$B^TB = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$
$$= \begin{bmatrix} \cos^2\theta + \operatorname{sen}^2\theta & -\cos\theta\operatorname{sen}\theta + \operatorname{sen}\theta\cos\theta \\ -\operatorname{sen}\theta\cos\theta + \cos\theta\operatorname{sen}\theta & \operatorname{sen}^2\theta + \cos^2\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Por tanto, B es ortogonal, por el Teorema 5.5, y

$$B^{-1} = B^T = \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ -\operatorname{sen}\theta & \cos\theta \end{bmatrix}$$



La palabra *isometría* significa literalmente "conservación de longitud", pues se deriva de las raíces griegas *isos* ("igual") y *metron* ("medida").

Comentario La matriz A del ejemplo 5.7 es un ejemplo de una matriz permutación, la cual se obtiene al permutar las columnas de una matriz identidad. En general, cualquier matriz permutación de $n \times n$ es ortogonal (vea el ejercicio 25). La matriz B es la matriz de una rotación en un ángulo θ en $\mathbb{R}^2$. Cualquier rotación tiene la propiedad de que es una transformación que *conserva la longitud* (que en geometría se conoce como ***isometría***). El siguiente teorema muestra que toda transformación con matriz ortogonal es una isometría. Las matrices ortogonales también conservan los productos punto. De hecho, las matrices ortogonales se caracterizan por alguna de dichas propiedades.

Teorema 5.6 Sea Q una matriz de $n \times n$. Los siguientes enunciados son equivalentes:

a. Q es ortogonal.
b. $\|Q\mathbf{x}\| = \|\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.
c. $Q\mathbf{x}\cdot Q\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$ para toda $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$.

Demostración Se demostrará que (a) ⇒ (c) ⇒ (b) ⇒ (a). Para hacerlo, será necesario usar el hecho de que, si $\mathbf{x}$ y $\mathbf{y}$ son vectores (columna) en $\mathbb{R}^n$, entonces $\mathbf{x}\cdot\mathbf{y} = \mathbf{x}^T\mathbf{y}$.

(a) ⇒ (c) Suponga que Q es ortogonal. Entonces $Q^TQ = I$ y se tiene

$$Q\mathbf{x}\cdot Q\mathbf{y} = (Q\mathbf{x})^TQ\mathbf{y} = \mathbf{x}^TQ^TQ\mathbf{y} = \mathbf{x}^TI\mathbf{y} = \mathbf{x}^T\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$$

(c) ⇒ (b) Suponga que $Q\mathbf{x}\cdot Q\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$ para toda $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$. Entonces, al tomar $\mathbf{y} = \mathbf{x}$, se tiene $Q\mathbf{x}\cdot Q\mathbf{x} = \mathbf{x}\cdot\mathbf{x}$, así $\|Q\mathbf{x}\| = \sqrt{Q\mathbf{x}\cdot Q\mathbf{x}} = \sqrt{\mathbf{x}\cdot\mathbf{x}} = \|\mathbf{x}\|$.

(b) ⇒ (a) Suponga que la propiedad (b) se cumple y sea $\mathbf{q}_i$ la i-ésima columna de Q. Al usar el ejercicio 63 en la sección 1.2 y la propiedad (b), se tiene

$$\begin{aligned} \mathbf{x}\cdot\mathbf{y} &= \tfrac{1}{4}(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2) \\ &= \tfrac{1}{4}(\|Q(\mathbf{x}+\mathbf{y})\|^2 - \|Q(\mathbf{x}-\mathbf{y})\|^2) \\ &= \tfrac{1}{4}(\|Q\mathbf{x}+Q\mathbf{y}\|^2 - \|Q\mathbf{x}-Q\mathbf{y}\|^2) \\ &= Q\mathbf{x}\cdot Q\mathbf{y} \end{aligned}$$

para toda $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$. [Esto demuestra que (b) ⇒ (c).]

Ahora, si $\mathbf{e}_i$ es el i-ésimo vector de la base estándar, entonces $\mathbf{q}_i = Q\mathbf{e}_i$. En consecuencia,

$$\mathbf{q}_i\cdot\mathbf{q}_j = Q\mathbf{e}_i\cdot Q\mathbf{e}_j = \mathbf{e}_i\cdot\mathbf{e}_j = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

Por tanto, las columnas de Q forman un conjunto ortonormal, de modo que Q es una matriz ortogonal.

Al observar las matrices ortogonales A y B en el ejemplo 5.7, podrá notar que no sólo sus columnas forman conjuntos ortonormales, también lo hacen los *renglones*. De hecho, toda matriz ortogonal tiene esta propiedad, como lo muestra el siguiente teorema.

Teorema 5.7 Si Q es una matriz ortogonal, entonces sus renglones forman un conjunto ortonormal.

Demostración Del Teorema 5.5 se sabe que $Q^{-1} = Q^T$. Por tanto,

$$(Q^T)^{-1} = (Q^{-1})^{-1} = Q = (Q^T)^T$$

de modo que Q^T es una matriz ortogonal. Por tanto, las columnas de Q^T, que son los renglones de Q, forman un conjunto ortonormal.

El teorema final en esta sección menciona algunas otras propiedades de las matrices ortogonales.

Teorema 5.8 Sea Q una matriz ortogonal.

a. Q^{-1} es ortogonal.
b. $\det Q = \pm 1$
c. Si λ es un eigenvalor de Q, entonces $|\lambda| = 1$.
d. Si Q_1 y Q_2 son matrices ortogonales de $n \times n$, entonces también lo es Q_1Q_2.

Demostración Se demostrará la propiedad (c) y las demostraciones de las propiedades restantes se dejarán como ejercicios.

(c) Sea λ un eigenvalor de Q con su correspondiente eigenvector $\mathbf{v}$. Entonces $Q\mathbf{v} = \lambda\mathbf{v}$ y, por el Teorema 5.6(b), se tiene

$$\|\mathbf{v}\| = \|Q\mathbf{v}\| = \|\lambda\mathbf{v}\| = |\lambda|\|\mathbf{v}\|$$

Dado que $\|\mathbf{v}\| \neq 0$, esto implica que $|\lambda| = 1$.

a + bi

Comentario La propiedad (c) se cumple incluso para eigenvalores complejos. La matriz $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ es ortogonal con eigenvalores i y $-i$, que tienen valor absoluto 1.

Ejercicios 5.1

En los ejercicios 1-6, determine cuáles conjuntos de vectores son ortogonales.

1. $\begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$

2. $\begin{bmatrix} 4 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -2 \\ 5 \\ 1 \end{bmatrix}$

3. $\begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 4 \end{bmatrix}$

4. $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} -2 \\ 6 \\ -1 \end{bmatrix}$

5. $\begin{bmatrix} 2 \\ 3 \\ -1 \\ 4 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ -6 \\ 2 \\ 7 \end{bmatrix}$

6. $\begin{bmatrix} 1 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 7-10, demuestre que los vectores dados forman una base ortogonal para $\mathbb{R}^2$ o $\mathbb{R}^3$. Luego use el Teorema 5.2 para expresar **w** *como una combinación lineal de dichos vectores base. Proporcione el vector coordenada $[\mathbf{w}]_{\mathcal{B}}$ de* **w** *con respecto a la base $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ de $\mathbb{R}^2$ o $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ de $\mathbb{R}^3$.*

7. $\mathbf{v}_1 = \begin{bmatrix} 4 \\ -2 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \mathbf{w} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$

8. $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -2 \\ 6 \end{bmatrix}; \mathbf{w} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

9. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}; \mathbf{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

10. $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ -1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}; \mathbf{w} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

En los ejercicios 11-15, determine si el conjunto ortogonal de vectores dado es ortonormal. Si no lo es, normalice los vectores para formar un conjunto ortonormal.

11. $\begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}, \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$

12. $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}$

13. $\begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ -\frac{5}{2} \end{bmatrix}$

14. $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \begin{bmatrix} 0 \\ \frac{1}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{6} \\ \frac{1}{6} \\ -\frac{1}{6} \end{bmatrix}$

15. $\begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 0 \\ \sqrt{6}/3 \\ 1/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}, \begin{bmatrix} \sqrt{3}/2 \\ -\sqrt{3}/6 \\ \sqrt{3}/6 \\ -\sqrt{3}/6 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$

En los ejercicios 16-21, determine si la matriz dada es ortogonal. Si lo es, encuentre su inversa.

16. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

17. $\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

18. $\begin{bmatrix} \frac{1}{2} & \frac{1}{3} & \frac{2}{5} \\ \frac{1}{2} & -\frac{1}{3} & \frac{2}{5} \\ -\frac{1}{2} & 0 & \frac{4}{5} \end{bmatrix}$

19. $\begin{bmatrix} \cos\theta \operatorname{sen}\theta & -\cos\theta & -\operatorname{sen}^2\theta \\ \cos^2\theta & \operatorname{sen}\theta & -\cos\theta \operatorname{sen}\theta \\ \operatorname{sen}\theta & 0 & \cos\theta \end{bmatrix}$

20. $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$

21. $\begin{bmatrix} 1 & 0 & 0 & 1/\sqrt{6} \\ 0 & 2/3 & 1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -2/3 & 1/\sqrt{2} & -1/\sqrt{6} \\ 0 & 1/3 & 0 & 1/\sqrt{2} \end{bmatrix}$

22. Demuestre el Teorema 5.8(a).

23. Demuestre el Teorema 5.8(b).

24. Demuestre el Teorema 5.8(d).

25. Demuestre que toda matriz permutación es ortogonal.

26. Si Q es una matriz ortogonal, demuestre que cualquier matriz obtenida al reordenar los renglones de Q también es ortogonal.

27. Sea Q una matriz ortogonal de 2×2 y sean **x** y **y** vectores en $\mathbb{R}^2$. Si θ es el ángulo entre **x** y **y**, demuestre que el ángulo entre $Q\mathbf{x}$ y $Q\mathbf{y}$ también es θ. (Esto demuestra que las transformaciones lineales definidas por matrices ortogonales *conservan el ángulo* en $\mathbb{R}^2$, un hecho que es verdadero en general.)

28. **(a)** Demuestre que una matriz ortogonal de 2×2 debe tener la forma

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} a & b \\ b & -a \end{bmatrix}$$

donde $\begin{bmatrix} a \\ b \end{bmatrix}$ es un vector unitario.

(b) Al usar el inciso (a), demuestre que toda matriz ortogonal de 2×2 es de la forma

$$\begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix} \quad \text{o} \quad \begin{bmatrix} \cos\theta & \operatorname{sen}\theta \\ \operatorname{sen}\theta & -\cos\theta \end{bmatrix}$$

donde $0 \le \theta < 2\pi$.

(c) Demuestre que toda matriz ortogonal de 2×2 corresponde a una rotación o a una reflexión en $\mathbb{R}^2$.

(d) Demuestre que una matriz Q ortogonal de 2×2 corresponde a una rotación en $\mathbb{R}^2$ si $\det Q = 1$ y a una reflexión en $\mathbb{R}^2$ si $\det Q = -1$.

En los ejercicios 29-32, use el ejercicio 28 para determinar si la matriz ortogonal dada representa una rotación o una reflexión. Si es una rotación, proporcione el ángulo de rotación; si es una reflexión, proporcione la recta de reflexión.

29. $\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

30. $\begin{bmatrix} -1/2 & \sqrt{3}/2 \\ -\sqrt{3}/2 & -1/2 \end{bmatrix}$

31. $\begin{bmatrix} -1/2 & \sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{bmatrix}$

32. $\begin{bmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{bmatrix}$

33. Sean A y B matrices ortogonales de $n \times n$.

(**a**) Demuestre que $A(A^T + B^T)B = A + B$.

(**b**) Use el inciso (a) para demostrar que, si $\det A + \det B = 0$, entonces $A + B$ no es invertible.

34. Sea $\mathbf{x}$ un vector unitario en $\mathbb{R}^n$. Particione $\mathbf{x}$ como

$$\mathbf{x} = \begin{bmatrix} x_1 \\ \hdashline x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ \mathbf{y} \end{bmatrix}$$

Sea

$$Q = \left[\begin{array}{c:c} x_1 & \mathbf{y}^T \\ \hdashline \mathbf{y} & I - \left(\dfrac{1}{1 - x_1}\right)\mathbf{y}\mathbf{y}^T \end{array}\right]$$

Demuestre que Q es ortogonal. (Este procedimiento proporciona un método rápido para encontrar una base ortonormal para $\mathbb{R}^n$ con un primer vector $\mathbf{x}$ prescrito, una construcción que frecuentemente es útil en aplicaciones.)

35. Demuestre que, si una matriz triangular superior es ortogonal, entonces debe ser una matriz diagonal.

36. Demuestre que, si $n > m$, entonces no hay una matriz A de $m \times n$ tal que $\|A\mathbf{x}\| = \|\mathbf{x}\|$ para toda $\mathbf{x}$ en $\mathbb{R}^n$.

37. Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base ortonormal para $\mathbb{R}^n$.

(**a**) Demuestre que, para cualquier $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$,

$$\mathbf{x}\cdot\mathbf{y} = (\mathbf{x}\cdot\mathbf{v}_1)(\mathbf{y}\cdot\mathbf{v}_1) + (\mathbf{x}\cdot\mathbf{v}_2)(\mathbf{y}\cdot\mathbf{v}_2) + \cdots + (\mathbf{x}\cdot\mathbf{v}_n)(\mathbf{y}\cdot\mathbf{v}_n)$$

(Esta identidad se llama ***Identidad de Parseval***.)

(**b**) ¿Qué implica la Identidad de Parseval acerca de la relación entre los productos punto $\mathbf{x}\cdot\mathbf{y}$ y $[\mathbf{x}]_{\mathcal{B}}\cdot[\mathbf{y}]_{\mathcal{B}}$?

5.2 Complementos y proyecciones ortogonales

En esta sección se generaliza dos conceptos que se encontraron en el capítulo 1. La noción de un vector normal a un plano se extenderá a complementos ortogonales y la proyección de un vector sobre otro dará lugar al concepto de proyección ortogonal sobre un subespacio.

$W^{\perp}$ se pronuncia "W perp".

Complementos ortogonales

Un vector normal $\mathbf{n}$ a un plano es ortogonal a todo vector en dicho plano. Si el plano pasa por el origen, entonces es un subespacio W de $\mathbb{R}^3$, como lo es gen($\mathbf{n}$). Por tanto, se tienen dos subespacios de $\mathbb{R}^3$ con la propiedad de que todo vector de uno es ortogonal a todo vector del otro. Esta es la idea detrás de la siguiente definición.

Definición Sea W un subespacio de $\mathbb{R}^n$. Se dice que un vector $\mathbf{v}$ en $\mathbb{R}^n$ es ***ortogonal a*** W si $\mathbf{v}$ es ortogonal a todo vector en W. El conjunto de todos los vectores que son ortogonales a W se llama ***complemento ortogonal de*** W y se denota $W^{\perp}$. Esto es,

$$W^{\perp} = \{\mathbf{v} \text{ en } \mathbb{R}^n : \mathbf{v}\cdot\mathbf{w} = 0 \quad \text{para todo } \mathbf{w} \text{ en } W\}$$

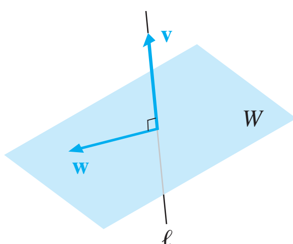


Figura 5.5
$\ell = W^{\perp}$ y $W = \ell^{\perp}$

Ejemplo 5.8

Si W es un plano que pasa por el origen en $\mathbb{R}^3$ y ℓ es la recta que pasa por el origen perpendicular a W (esto es, paralela al vector normal a W), entonces todo vector $\mathbf{v}$ en ℓ es ortogonal a todo vector $\mathbf{w}$ en W; en consecuencia, $\ell = W^{\perp}$. Más aún, W consiste *precisamente* de aquellos vectores $\mathbf{w}$ que son ortogonales a todo $\mathbf{v}$ en ℓ; por tanto, también se tiene $W = \ell^{\perp}$. La figura 5.5 ilustra esta situación.

En el ejemplo 5.8, el complemento ortogonal de un subespacio evidencia ser otro subespacio. Además, el complemento del complemento de un subespacio es el subespacio original. Estas propiedades son verdaderas en general y se prueban como las propiedades (a) y (b) del Teorema 5.9. Las propiedades (c) y (d) también serán útiles. (Recuerde que la *intersección* $A \cap B$ de los conjuntos A y B consiste de sus elementos comunes. Vea el Apéndice A.)

Teorema 5.9 Sea W un subespacio de $\mathbb{R}^n$.

a. $W^\perp$ es un subespacio de $\mathbb{R}^n$.
b. $(W^\perp)^\perp = W$
c. $W \cap W^\perp = \{\mathbf{0}\}$
d. Si $W = \text{gen}(\mathbf{w}_1, \ldots, \mathbf{w}_k)$, entonces $\mathbf{v}$ está en $W^\perp$ si y sólo si $\mathbf{v} \cdot \mathbf{w}_i = 0$ para todo $i = 1, \ldots, k$.

Demostración (a) Dado que $\mathbf{0} \cdot \mathbf{w} = 0$ para todo $\mathbf{w}$ en W, $\mathbf{0}$ está en $W^\perp$. Sean $\mathbf{u}$ y $\mathbf{v}$ que están en $W^\perp$ y sea c un escalar. Entonces

$$\mathbf{u} \cdot \mathbf{w} = \mathbf{v} \cdot \mathbf{w} = 0 \quad \text{para todo } \mathbf{w} \text{ en } W$$

Por tanto,

$$(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w} = 0 + 0 = 0$$

de modo que $\mathbf{u} + \mathbf{v}$ está en $W^\perp$.

También se tiene

$$(c\mathbf{u}) \cdot \mathbf{w} = c(\mathbf{u} \cdot \mathbf{w}) = c(0) = 0$$

de donde se ve que $c\mathbf{u}$ está en $W^\perp$. Se concluye que $W^\perp$ es un subespacio de $\mathbb{R}^n$.

(b) Esta propiedad se demostrará como el Corolario 5.12.

(c) Se le pide demostrar esta propiedad en el ejercicio 23.

(d) Se le pide demostrar esta propiedad en el ejercicio 24.

Ahora puede expresar algunas relaciones fundamentales que involucran los subespacios asociados con una matriz de $m \times n$.

Teorema 5.10 Sea A una matriz de $m \times n$. Entonces el complemento ortogonal del espacio renglón de A es el espacio nulo de A y el complemento ortogonal del espacio columna de A es el espacio nulo de A^T:

$$(\text{renglón}(A))^\perp = \text{nulo}(A) \quad \text{y} \quad (\text{col}(A))^\perp = \text{nulo}(A^T)$$

Demostración Si $\mathbf{x}$ es un vector en $\mathbb{R}^n$, entonces $\mathbf{x}$ está en $(\text{renglón}(A))^\perp$ si y sólo si $\mathbf{x}$ es ortogonal a todo renglón de A. Pero esto es verdadero si y sólo si $A\mathbf{x} = \mathbf{0}$, lo cual equivale a que $\mathbf{x}$ está en nulo(A), de modo que se ha establecido la primera identidad. Para demostrar la segunda identidad, simplemente sustituya A por A^T y use el hecho de que renglón$(A^T) = \text{col}(A)$.

Por tanto, una matriz de $m \times n$ tiene cuatro subespacios: renglón(A), nulo(A), col(A) y nulo(A^T). Los primeros dos son complementos ortogonales en $\mathbb{R}^n$ y los últimos

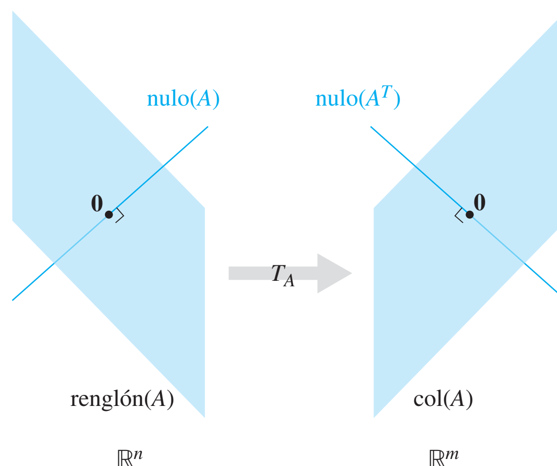


Figura 5.6
Los cuatro subespacios fundamentales

dos son complementos ortogonales en $\mathbb{R}^m$. La matriz A $m \times n$ define una transformación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ cuyo rango es col(A). Más aún, esta transformación envía nulo(A) a $\mathbf{0}$ en $\mathbb{R}^m$. La figura 5.6 ilustra esquemáticamente estas ideas. Estos cuatro subespacios se llaman los ***subespacios fundamentales*** de la matriz A de $m \times n$.

Ejemplo 5.9 Encuentre bases para los cuatro subespacios fundamentales de

$$A = \begin{bmatrix} 1 & 1 & 3 & 1 & 6 \\ 2 & -1 & 0 & 1 & -1 \\ -3 & 2 & 1 & -2 & 1 \\ 4 & 1 & 6 & 1 & 3 \end{bmatrix}$$

y verifique el Teorema 5.10.

Solución En los ejemplos 3.45, 3.47 y 3.48, calcularon las bases para el espacio renglón, espacio columna y espacio nulo de A. Se encontró que renglón$(A) = \text{gen}(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$, donde

$$\mathbf{u}_1 = [1 \quad 0 \quad 1 \quad 0 \quad -1], \quad \mathbf{u}_2 = [0 \quad 1 \quad 2 \quad 0 \quad 3], \quad \mathbf{u}_3 = [0 \quad 0 \quad 0 \quad 1 \quad 4]$$

Además, nulo$(A) = \text{gen}(\mathbf{x}_1, \mathbf{x}_2)$, donde

$$\mathbf{x}_1 = \begin{bmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ -3 \\ 0 \\ -4 \\ 1 \end{bmatrix}$$

Para demostrar $(\text{renglón}(A))^\perp = \text{nulo}(A)$, es suficiente demostrar que todo $\mathbf{u}_i$ es ortogonal a cada $\mathbf{x}_j$, lo cual es un ejercicio sencillo. (¿Por qué esto es suficiente?)

El espacio columna de A es $\text{col}(A) = \text{gen}(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$, donde

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 2 \\ -3 \\ 4 \end{bmatrix}, \quad \mathbf{a}_2 = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix} 1 \\ 1 \\ -2 \\ 1 \end{bmatrix}$$

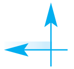
Todavía es necesario calcular el espacio nulo de A^T. La reducción por renglones produce

$$[A^T \mid \mathbf{0}] = \left[\begin{array}{cccc|c} 1 & 2 & -3 & 4 & 0 \\ 1 & -1 & 2 & 1 & 0 \\ 3 & 0 & 1 & 6 & 0 \\ 1 & 1 & -2 & 1 & 0 \\ 6 & -1 & 1 & 3 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 6 & 0 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

De este modo, si $\mathbf{y}$ está en el espacio nulo de A^T, entonces $y_1 = -y_4$, $y_2 = -6y_4$ y $y_3 = -3y_4$. Se tiene que

$$\text{nulo}(A^T) = \left\{ \begin{bmatrix} -y_4 \\ -6y_4 \\ -3y_4 \\ y_4 \end{bmatrix} \right\} = \text{gen}\left(\begin{bmatrix} -1 \\ -6 \\ -3 \\ 1 \end{bmatrix} \right)$$

y es fácil comprobar que este vector es ortogonal a $\mathbf{a}_1$, $\mathbf{a}_2$ y $\mathbf{a}_3$.



El método del ejemplo 5.9 se adapta con facilidad a otras situaciones.

Ejemplo 5.10

Sea W el subespacio de $\mathbb{R}^5$ generado por

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ -3 \\ 5 \\ 0 \\ 5 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \\ -2 \\ 3 \end{bmatrix}, \quad \mathbf{w}_3 = \begin{bmatrix} 0 \\ -1 \\ 4 \\ -1 \\ 5 \end{bmatrix}$$

Encuentre una base para $W^\perp$.

Solución El subespacio W generado por $\mathbf{w}_1$, $\mathbf{w}_2$ y $\mathbf{w}_3$ es lo mismo que el espacio columna de

$$A = \begin{bmatrix} 1 & -1 & 0 \\ -3 & 1 & -1 \\ 5 & 2 & 4 \\ 0 & -2 & -1 \\ 5 & 3 & 5 \end{bmatrix}$$

En consecuencia, por el Teorema 5.10, $W^\perp = (\text{col}(A))^\perp = \text{nulo}(A^T)$ y puede proceder como en el ejemplo anterior. Calcule

$$[A^T \mid \mathbf{0}] = \left[\begin{array}{rrrrr|r} 1 & -3 & 5 & 0 & 5 & 0 \\ -1 & 1 & 2 & -2 & 3 & 0 \\ 0 & -1 & 4 & -1 & 5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{rrrrr|r} 1 & 0 & 0 & 3 & 4 & 0 \\ 0 & 1 & 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \end{array}\right]$$

Por tanto, **y** está en $W^\perp$ si y sólo si $y_1 = -3y_4 - 4y_5$, $y_2 = -y_4 - 3y_5$, y $y_3 = -2y_5$. Se tiene que

$$W^\perp = \left\{ \begin{bmatrix} -3y_4 - 4y_5 \\ -y_4 - 3y_5 \\ -2y_5 \\ y_4 \\ y_5 \end{bmatrix} \right\} = \text{gen}\left(\begin{bmatrix} -3 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ -3 \\ -2 \\ 0 \\ 1 \end{bmatrix} \right)$$

y estos dos vectores forman una base para $W^\perp$.

Proyecciones ortogonales

Recuerde que, en $\mathbb{R}^2$, la proyección de un vector **v** sobre un vector distinto de cero **u** está dada por

$$\text{proy}_{\mathbf{u}}(\mathbf{v}) = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u}$$

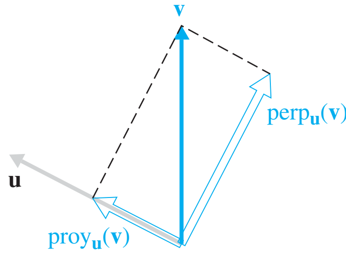


Figura 5.7
$\mathbf{v} = \text{proy}_{\mathbf{u}}(\mathbf{v}) + \text{perp}_{\mathbf{u}}(\mathbf{v})$

Más aún, el vector $\text{perp}_{\mathbf{u}}(\mathbf{v}) = \mathbf{v} - \text{proy}_{\mathbf{u}}(\mathbf{v})$ es ortogonal a $\text{proy}_{\mathbf{u}}(\mathbf{v})$, y **v** puede descomponerse como

$$\mathbf{v} = \text{proy}_{\mathbf{u}}(\mathbf{v}) + \text{perp}_{\mathbf{u}}(\mathbf{v})$$

como se muestra en la figura 5.7.

Si se hace $W = \text{gen}(\mathbf{u})$, entonces $\mathbf{w} = \text{proy}_{\mathbf{u}}(\mathbf{v})$ está en W y $\mathbf{w}^\perp = \text{perp}_{\mathbf{u}}(\mathbf{v})$ está en $W^\perp$. En consecuencia, se tiene una forma de "descomponer" **v** en la suma de dos vectores, uno de W y el otro ortogonal a W; a saber, $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$. Ahora esta idea se generaliza a $\mathbb{R}^n$.

Definición Sea W un subespacio de $\mathbb{R}^n$ y sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ una base ortogonal para W. Para cualquier vector **v** en $\mathbb{R}^n$, la ***proyección ortogonal de v sobre W*** se define como

$$\text{proy}_W(\mathbf{v}) = \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \cdots + \left(\frac{\mathbf{u}_k \cdot \mathbf{v}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)\mathbf{u}_k$$

El ***componente ortogonal de v a W*** es el vector

$$\text{perp}_W(\mathbf{v}) = \mathbf{v} - \text{proy}_W(\mathbf{v})$$

Cada sumando en la definición de $\text{proy}_W(\mathbf{v})$ también es una proyección sobre un solo vector (o, de manera equivalente, el subespacio unidimensional generado por él; en el sentido anterior). Por tanto, con la notación de la definición anterior, puede escribir

$$\text{proy}_W(\mathbf{v}) = \text{proy}_{\mathbf{u}_1}(\mathbf{v}) + \cdots + \text{proy}_{\mathbf{u}_k}(\mathbf{v})$$

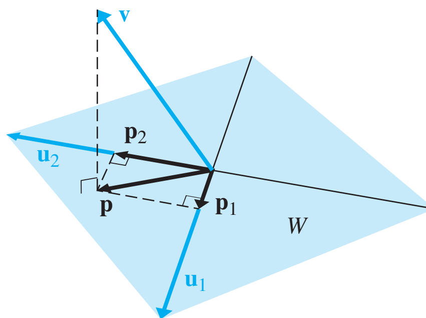


Figura 5.8

$\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2$

Dado que los vectores $\mathbf{u}_i$ son ortogonales, la proyección ortogonal de $\mathbf{v}$ sobre W es la suma de sus proyecciones sobre subespacios unidimensionales que son mutuamente ortogonales. La figura 5.8 ilustra esta situación con $W = \text{gen}(\mathbf{u}_1, \mathbf{u}_2)$, $\mathbf{p} = \text{proy}_W(\mathbf{v})$, $\mathbf{p}_1 = \text{proy}_{\mathbf{u}_1}(\mathbf{v})$ y $\mathbf{p}_2 = \text{proy}_{\mathbf{u}_2}(\mathbf{v})$.

Como un caso especial de la definición de $\text{proy}_W(\mathbf{v})$, ahora también se tiene una amigable interpretación geométrica del Teorema 5.2. En términos de la notación y terminología actuales, dicho teorema afirma que si $\mathbf{w}$ está en el subespacio W de $\mathbb{R}^n$, que tiene base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$, entonces

$$\begin{aligned}\mathbf{w} &= \left(\frac{\mathbf{w}\cdot\mathbf{v}_1}{\mathbf{v}_1\cdot\mathbf{v}_1}\right)\mathbf{v}_1 + \cdots + \left(\frac{\mathbf{w}\cdot\mathbf{v}_k}{\mathbf{v}_k\cdot\mathbf{v}_k}\right)\mathbf{v}_k \\ &= \text{proy}_{\mathbf{v}_1}(\mathbf{w}) + \cdots + \text{proy}_{\mathbf{v}_k}(\mathbf{w})\end{aligned}$$

Por tanto, $\mathbf{w}$ se descompone en una suma de proyecciones ortogonales sobre subespacios unidimensionales mutuamente ortogonales de W.

La definición anterior parece depender de la elección de la base ortogonal; esto es, una base diferente $\{\mathbf{u}'_1, \ldots, \mathbf{u}'_k\}$ para W parecería producir "diferentes" $\text{proy}_W(\mathbf{v})$ y $\text{perp}_W(\mathbf{v})$. Por fortuna, este no es el caso, como pronto se probará. Por ahora, confórmese con un ejemplo.

Ejemplo 5.11

Sea W el plano en $\mathbb{R}^3$ con ecuación $x - y + 2z = 0$ y sea $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$. Encuentre la proyección ortogonal de $\mathbf{v}$ sobre W y el componente ortogonal de $\mathbf{v}$ a W.

Solución En el ejemplo 5.3 se encontró una base ortogonal para W. Considere

$$\mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

se tiene

$$\begin{aligned} \mathbf{u}_1\cdot\mathbf{v} &= 2 \qquad & \mathbf{u}_2\cdot\mathbf{v} &= -2 \\ \mathbf{u}_1\cdot\mathbf{u}_1 &= 2 \qquad & \mathbf{u}_2\cdot\mathbf{u}_2 &= 3 \end{aligned}$$

En consecuencia,

$$\text{proy}_W(\mathbf{v}) = \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{u}_2 \cdot \mathbf{v}}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right)\mathbf{u}_2$$
$$= \tfrac{2}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \tfrac{2}{3}\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

y
$$\text{perp}_W(\mathbf{v}) = \mathbf{v} - \text{proy}_W(\mathbf{v}) = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} = \begin{bmatrix} \frac{4}{3} \\ -\frac{4}{3} \\ \frac{8}{3} \end{bmatrix}$$

Es fácil ver que $\text{proy}_W(\mathbf{v})$ está en W, pues satisface la ecuación del plano. Igualmente, es fácil ver que $\text{perp}_W(\mathbf{v})$ es ortogonal a W, pues es un múltiplo escalar del vector normal $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ a W. (Vea la figura 5.9.)

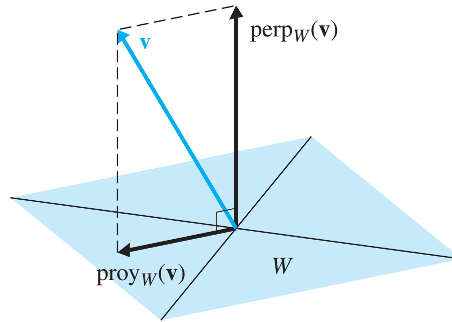


Figura 5.9
$\mathbf{v} = \text{proy}_W(\mathbf{v}) + \text{perp}_W(\mathbf{v})$

El siguiente teorema muestra que siempre es posible encontrar una descomposición de un vector con respecto a un subespacio y su complemento ortogonal.

Teorema 5.11 **El teorema de descomposición ortogonal**

Sea W un subespacio de $\mathbb{R}^n$ y sea $\mathbf{v}$ un vector en $\mathbb{R}^n$. Entonces existen vectores únicos $\mathbf{w}$ en W y $\mathbf{w}^\perp$ en $W^\perp$ tales que

$$\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$$

Demostración Es necesario demostrar dos cosas: que tal descomposición *existe* y que es *única*.

Para demostrar su existencia se elige una base ortogonal $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ para W. Sea $\mathbf{w} = \text{proy}_W(\mathbf{v})$ y sea $\mathbf{w}^\perp = \text{perp}_W(\mathbf{v})$. Entonces

$$\mathbf{w} + \mathbf{w}^\perp = \text{proy}_W(\mathbf{v}) + \text{perp}_W(\mathbf{v}) = \text{proy}_W(\mathbf{v}) + (\mathbf{v} - \text{proy}_W(\mathbf{v})) = \mathbf{v}$$

Claramente, $\mathbf{w} = \text{proy}_W(\mathbf{v})$ está en W, pues es una combinación lineal de los vectores base $\mathbf{u}_1, \ldots, \mathbf{u}_k$. Para demostrar que $\mathbf{w}^\perp$ está en $W^\perp$, es suficiente demostrar que $\mathbf{w}^\perp$ es ortogonal a cada uno de los vectores base $\mathbf{u}_i$, por el Teorema 5.9(d). Calcule

$$\begin{aligned}
\mathbf{u}_i \cdot \mathbf{w}^\perp &= \mathbf{u}_i \cdot \text{perp}_W(\mathbf{v}) \\
&= \mathbf{u}_i \cdot (\mathbf{v} - \text{proy}_W(\mathbf{v})) \\
&= \mathbf{u}_i \cdot \left(\mathbf{v} - \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 - \cdots - \left(\frac{\mathbf{u}_k \cdot \mathbf{v}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)\mathbf{u}_k\right) \\
&= \mathbf{u}_i \cdot \mathbf{v} - \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)(\mathbf{u}_i \cdot \mathbf{u}_1) - \cdots - \left(\frac{\mathbf{u}_i \cdot \mathbf{v}}{\mathbf{u}_i \cdot \mathbf{u}_i}\right)(\mathbf{u}_i \cdot \mathbf{u}_i) - \cdots \\
&\quad - \left(\frac{\mathbf{u}_k \cdot \mathbf{v}}{\mathbf{u}_k \cdot \mathbf{u}_k}\right)(\mathbf{u}_i \cdot \mathbf{u}_k) \\
&= \mathbf{u}_i \cdot \mathbf{v} - 0 - \cdots - \left(\frac{\mathbf{u}_i \cdot \mathbf{v}}{\mathbf{u}_i \cdot \mathbf{u}_i}\right)(\mathbf{u}_i \cdot \mathbf{u}_i) - \cdots - 0 \\
&= \mathbf{u}_i \cdot \mathbf{v} - \mathbf{u}_i \cdot \mathbf{v} = 0
\end{aligned}$$

pues $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ para $j \neq i$. Esto prueba que $\mathbf{w}^\perp$ está en $W^\perp$ y completa la parte de la demostración relativa a la existencia.

Para demostrar la unicidad de esta descomposición, suponga que tiene otra descomposición $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_1^\perp$, donde $\mathbf{w}_1$ está en W y $\mathbf{w}_1^\perp$ está en $W^\perp$. Entonces $\mathbf{w} + \mathbf{w}^\perp = \mathbf{w}_1 + \mathbf{w}_1^\perp$, de modo que

$$\mathbf{w} - \mathbf{w}_1 = \mathbf{w}_1^\perp - \mathbf{w}^\perp$$

Pero, dado que $\mathbf{w} - \mathbf{w}_1$ está en W y $\mathbf{w}_1^\perp - \mathbf{w}^\perp$ está en $W^\perp$ (porque son subespacios), se sabe que este vector común está en $W \cap W^\perp = \{\mathbf{0}\}$ [por el Teorema 5.9(c)]. Por ende,

$$\mathbf{w} - \mathbf{w}_1 = \mathbf{w}_1^\perp - \mathbf{w}^\perp = 0$$

así que $\mathbf{w}_1 = \mathbf{w}$ y $\mathbf{w}_1^\perp = \mathbf{w}_1$.

El ejemplo 5.11 ilustró el teorema de descomposición ortogonal. Cuando W es el subespacio de $\mathbb{R}^3$ dado por el plano con ecuación $x - y + 2z = 0$, la descomposición ortogonal de $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$ con respecto a W es $\mathbf{v} = \mathbf{w} + \mathbf{w}^\perp$, donde

$$\mathbf{w} = \text{proy}_W(\mathbf{v}) = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix} \quad \text{y} \quad \mathbf{w}^\perp = \text{perp}_W(\mathbf{v}) = \begin{bmatrix} \frac{4}{3} \\ -\frac{4}{3} \\ \frac{8}{3} \end{bmatrix}$$

La unicidad de la descomposición ortogonal garantiza que las definiciones de $\text{proy}_W(\mathbf{v})$ y $\text{perp}_W(\mathbf{v})$ no dependan de la elección de la base ortogonal. El teorema de descomposición ortogonal también permite probar la propiedad (b) del Teorema 5.9. Aquí, dicha propiedad se enuncia como un corolario al teorema de descomposición ortogonal.

Corolario 5.12 Si W es un subespacio de $\mathbb{R}^n$, entonces

$$(W^\perp)^\perp = W$$

Demostración Si $\mathbf{w}$ está en W y $\mathbf{x}$ está en $W^{\perp}$, entonces $\mathbf{w} \cdot \mathbf{x} = 0$. Pero ahora esto implica que $\mathbf{w}$ está en $(W^{\perp})^{\perp}$. En consecuencia, $W \subseteq (W^{\perp})^{\perp}$. Para ver que en realidad aquí se tiene una igualdad, suponga lo contrario. Entonces existe un vector $\mathbf{v}$ en $(W^{\perp})^{\perp}$ que no está en W. Por el Teorema 5.11, se puede escribir $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$ para vectores (únicos) $\mathbf{w}$ en W y $\mathbf{w}^{\perp}$ en $W^{\perp}$. Pero ahora

$$0 = \mathbf{v} \cdot \mathbf{w}^{\perp} = (\mathbf{w} + \mathbf{w}^{\perp}) \cdot \mathbf{w}^{\perp} = \mathbf{w} \cdot \mathbf{w}^{\perp} + \mathbf{w}^{\perp} \cdot \mathbf{w}^{\perp} = 0 + \mathbf{w}^{\perp} \cdot \mathbf{w}^{\perp} = \mathbf{w}^{\perp} \cdot \mathbf{w}^{\perp}$$

de modo que $\mathbf{w}^{\perp} = \mathbf{0}$. Por tanto, $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp} = \mathbf{w}$ y en consecuencia $\mathbf{v}$ *está* en W, lo que es una contradicción. Se concluye que $(W^{\perp})^{\perp} = W$, como se requería.

También existe una buena relación entre las dimensiones de W y $W^{\perp}$, expresada en el Teorema 5.13.

Teorema 5.13 Si W es un subespacio de $\mathbb{R}^n$, entonces

$$\dim W + \dim W^{\perp} = n$$

Demostración Sea $\{\mathbf{u}_1, \ldots, \dim \mathbf{u}_k\}$ una base ortogonal para W y sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_l\}$ una base ortogonal para $W^{\perp}$. Entonces $\dim W = k$ y $\dim W^{\perp} = l$. Sea $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_k, \mathbf{v}_1, \ldots, \mathbf{v}_l\}$. Se afirma que $\mathcal{B}$ es una base ortogonal para $\mathbb{R}^n$.

Primero note que, dado que cada $\mathbf{u}_i$ está en W y cada $\mathbf{v}_j$ está en $W^{\perp}$,

$$\mathbf{u}_i \cdot \mathbf{v}_j = 0 \quad \text{para} \quad i = 1, \ldots, k \text{ y } j = 1, \ldots, l$$

Por tanto, $\mathcal{B}$ es un conjunto ortogonal y, en consecuencia, es linealmente independiente, por el Teorema 5.1. A continuación, si $\mathbf{v}$ es un vector en $\mathbb{R}^n$, el teorema de descomposición ortogonal dice que $\mathbf{v} = \mathbf{w} + \mathbf{w}^{\perp}$ para algún $\mathbf{w}$ en W y $\mathbf{w}^{\perp}$ en $W^{\perp}$. Dado que $\mathbf{w}$ puede escribirse como una combinación lineal de los vectores $\mathbf{u}_i$ y $\mathbf{w}^{\perp}$ puede escribirse como una combinación lineal de los vectores $\mathbf{v}_j$, $\mathbf{v}$ puede escribirse como una combinación lineal de los vectores en $\mathcal{B}$. Por tanto, $\mathcal{B}$ también genera a $\mathbb{R}^n$ y por tanto es una base para $\mathbb{R}^n$. Se tiene que $k + l = \dim \mathbb{R}^n$, o

$$\dim W + \dim W^{\perp} = n$$

Como bono adorable, cuando este resultado se aplica a los subespacios fundamentales de una matriz, se obtiene una demostración rápida del teorema del rank (Teorema 3.26), parafraseado aquí como Corolario 5.14.

Corolario 5.14 **El teorema del rank**

Si A es una matriz de $m \times n$, entonces

$$\operatorname{rank}(A) + \operatorname{nulidad}(A) = n$$

Demostración En el Teorema 5.13, tome $W = \text{renglón}(A)$. Entonces $W^{\perp} = \text{nulo}(A)$, por el Teorema 5.10, de modo que $\dim W = \operatorname{rank}(A)$ y $\dim W^{\perp} = \operatorname{nulidad}(A)$. El resultado viene a continuación.

Note que se obtiene una identidad contrapuesta al tomar $W = \operatorname{col}(A)$ [y por tanto $W^{\perp} = \text{nulo}(A^T)$]:

$$\operatorname{rank}(A) + \operatorname{nulidad}(A^T) = m$$

Las secciones 5.1 y 5.2 ilustraron algunas de las ventajas de trabajar con bases ortogonales. Sin embargo, no se estableció que todo subespacio *tiene* una base ortogonal, ni se proporcionó un método para construir tal base (excepto en casos particulares, como el ejemplo 5.3). Estos temas son la materia de la siguiente sección.

Ejercicios 5.2

En los ejercicios 1-6, encuentre el complemento ortogonal $W^\perp$ de W y proporcione una base para $W^\perp$.

1. $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 2x - y = 0 \right\}$

2. $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : 3x + 2y = 0 \right\}$

3. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + y - z = 0 \right\}$

4. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : -x + 3y - 5z = 0 \right\}$

5. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x = t, y = -t, z = 3t \right\}$

6. $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x = 2t, y = -t, z = \frac{1}{2}t \right\}$

En los ejercicios 7 y 8, encuentre bases para el espacio renglón y el espacio nulo de A. Verifique que todo vector en renglón(A) es ortogonal a todo vector en nulo(A).

7. $A = \begin{bmatrix} 1 & -1 & 3 \\ 5 & 2 & 1 \\ 0 & 1 & -2 \\ -1 & -1 & 1 \end{bmatrix}$

8. $A = \begin{bmatrix} 1 & 1 & 0 & 1 & -1 \\ -2 & 0 & 4 & 0 & 2 \\ 2 & 2 & -2 & 3 & 1 \\ 2 & 4 & 2 & 5 & 1 \end{bmatrix}$

En los ejercicios 9 y 10, encuentre bases para el espacio columna de A y el espacio nulo de A^T para el ejercicio dado. Verifique que todo vector en col(A) es ortogonal a todo vector en nulo(A^T).

9. Ejercicio 7 **10.** Ejercicio 8

En los ejercicios 11-14, sea W el subespacio generado por los vectores dados. Encuentre una base para $W^\perp$.

11. $\mathbf{w}_1 = \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$

12. $\mathbf{w}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix}$

13. $\mathbf{w}_1 = \begin{bmatrix} 2 \\ -1 \\ 6 \\ 3 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} -1 \\ 2 \\ -3 \\ -2 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 2 \\ 5 \\ 6 \\ 1 \end{bmatrix}$

14. $\mathbf{w}_1 = \begin{bmatrix} 3 \\ 2 \\ 0 \\ -1 \\ 4 \end{bmatrix}, \mathbf{w}_2 = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 0 \\ 1 \end{bmatrix}, \mathbf{w}_3 = \begin{bmatrix} 3 \\ -2 \\ 6 \\ -2 \\ 5 \end{bmatrix}$

En los ejercicios 15-18, encuentre la proyección ortogonal de $\mathbf{v}$ *sobre el subespacio W generado por los vectores* $\mathbf{u}_i$. *(Puede suponer que los vectores* $\mathbf{u}_i$ *son ortogonales.)*

15. $\mathbf{v} = \begin{bmatrix} 7 \\ -4 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

16. $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

17. $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix}$

18. $\mathbf{v} = \begin{bmatrix} 4 \\ -2 \\ -3 \\ 2 \end{bmatrix}, \mathbf{u}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{u}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$

En los ejercicios 19-22, encuentre la descomposición ortogonal de $\mathbf{v}$ *con respecto a W.*

19. $\mathbf{v} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, W = \text{gen}\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right)$

20. $\mathbf{v} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}, W = \text{gen}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}\right)$

21. $\mathbf{v} = \begin{bmatrix} 4 \\ -2 \\ 3 \end{bmatrix}, W = \text{gen}\left(\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}\right)$

22. $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \\ 5 \\ 3 \end{bmatrix}, W = \text{gen}\left(\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}\right)$

23. Demuestre el Teorema 5.9(c).

24. Demuestre el Teorema 5.9(d).

25. Sea W un subespacio de $\mathbb{R}^n$ y $\mathbf{v}$ un vector en $\mathbb{R}^n$. Suponga que $\mathbf{w}$ y $\mathbf{w}'$ son vectores ortogonales con $\mathbf{w}$ en W y que $\mathbf{v} = \mathbf{w} + \mathbf{w}'$. ¿Es necesariamente cierto que $\mathbf{w}'$ está en $W^{\perp}$? Demuestre que es verdadero o encuentre un contraejemplo.

26. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base ortogonal para $\mathbb{R}^n$ y sea $W = \text{gen}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$. ¿Es necesariamente cierto que $W^{\perp} = \text{gen}(\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$? Demuestre que es verdadero o encuentre un contraejemplo.

En los ejercicios 27-29, sea W un subespacio de $\mathbb{R}^n$ *y sea* $\mathbf{x}$ *un vector en* $\mathbb{R}^n$.

27. Demuestre que $\mathbf{x}$ está en W si y sólo si $\text{proy}_W(\mathbf{x}) = \mathbf{x}$.

28. Demuestre que $\mathbf{x}$ es ortogonal a W si y sólo si $\text{proy}_W(\mathbf{x}) = \mathbf{0}$.

29. Demuestre que $\text{proy}_W(\text{proy}_W(\mathbf{x})) = \text{proy}_W(\mathbf{x})$.

30. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ un conjunto ortogonal en $\mathbb{R}^n$ y sea $\mathbf{x}$ un vector en $\mathbb{R}^n$.

(a) Demuestre que

$$\|\mathbf{x}\|^2 \geq |\mathbf{x}\cdot\mathbf{v}_1|^2 + |\mathbf{x}\cdot\mathbf{v}_2|^2 + \cdots + |\mathbf{x}\cdot\mathbf{v}_k|^2$$

(Esta desigualdad se llama ***desigualdad de Bessel***.)

(b) Demuestre que la desigualdad de Bessel es una igualdad si y sólo si $\mathbf{x}$ está en gen(S).

5.3 El proceso de Gram-Schmidt y la factorización *QR*

En esta sección se presenta un método simple para construir una base ortogonal (u ortonormal) para cualquier subespacio de $\mathbb{R}^n$. Este método conducirá entonces a una de las más útiles de todas las factorizaciones de matrices.

El proceso de Gram-Schmidt

Se pide encontrar una base ortogonal para un subespacio W de $\mathbb{R}^n$. La idea es comenzar con una base arbitraria $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ para W y "ortogonalizarla" de vector en vector. La construcción básica se ilustrará con el subespacio W del ejemplo 5.3.

Ejemplo 5.12

Sea $W = \text{gen}(\mathbf{x}_1, \mathbf{x}_2)$, donde

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{x}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$

Construya una base ortogonal para W.

Solución A partir de $\mathbf{x}_1$, se obtiene un segundo vector que es ortogonal a él al tomar el componente ortogonal de $\mathbf{x}_2$ a $\mathbf{x}_1$ (figura 5.10).

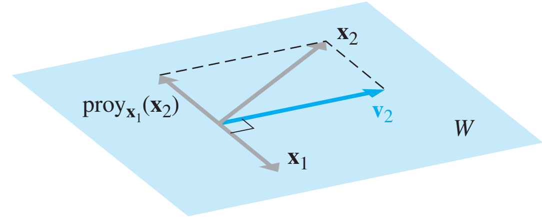
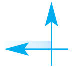


Figura 5.10
Construcción de $\mathbf{v}_2$ ortogonal a $\mathbf{x}_1$

Algebraicamente, se establece $\mathbf{v}_1 = \mathbf{x}_1$, de modo que

$$\begin{aligned}\mathbf{v}_2 &= \text{perp}_{\mathbf{x}_1}(\mathbf{x}_2) = \mathbf{x}_2 - \text{proy}_{\mathbf{x}_1}(\mathbf{x}_2)\\ &= \mathbf{x}_2 - \left(\frac{\mathbf{x}_1 \cdot \mathbf{x}_2}{\mathbf{x}_1 \cdot \mathbf{x}_1}\right)\mathbf{x}_1\\ &= \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - \left(\frac{-2}{2}\right)\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}\end{aligned}$$

Entonces $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto ortogonal de vectores en W. Por tanto, $\{\mathbf{v}_1, \mathbf{v}_2\}$ es un conjunto linealmente independiente y en consecuencia una base para W, pues $\dim W = 2$.



Comentario Observe que este método depende del *orden* de los vectores base originales. En el ejemplo 5.12, si hubiera tomado $\mathbf{x}_1 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$ y $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, habría obtenido



una base ortogonal diferente para W. (Verifique esto.)

La generalización de este método a más de dos vectores comienza como en el ejemplo 5.12. Entonces el proceso es construir de manera iterativa los componentes ortogonales de los vectores subsecuentes a todos los vectores que ya se construyeron. El método se conoce como ***proceso de Gram-Schmidt***.

Teorema 5.15 El proceso de Gram-Schmidt

Sea $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ una base para un subespacio W de $\mathbb{R}^n$ y defina lo siguiente:

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1, & W_1 &= \text{gen}(\mathbf{x}_1)\\
\mathbf{v}_2 &= \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_2}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1, & W_2 &= \text{gen}(\mathbf{x}_1, \mathbf{x}_2)\\
\mathbf{v}_3 &= \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{x}_3}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2, & W_3 &= \text{gen}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)\\
&\ \vdots\\
\mathbf{v}_k &= \mathbf{x}_k - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_k}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{x}_k}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 - \cdots\\
&\quad - \left(\frac{\mathbf{v}_{k-1} \cdot \mathbf{x}_k}{\mathbf{v}_{k-1} \cdot \mathbf{v}_{k-1}}\right)\mathbf{v}_{k-1}, & W_k &= \text{gen}(\mathbf{x}_1, \ldots, \mathbf{x}_k)
\end{aligned}$$

Entonces, para cada $i = 1, \ldots, k$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ es una base ortogonal para W_i. En particular, $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es una base ortogonal para W.

Jörgen Pedersen Gram (1850-1916) fue un actuario (estadístico de seguros) danés que estaba interesado en la ciencia de la medición. Primero publicó el proceso que lleva su nombre en un ensayo de 1883 acerca de mínimos cuadrados.
Erhard Schmidt (1876-1959) fue un matemático alemán discípulo del gran David Hilbert y es considerado uno de los fundadores de la rama de las matemáticas conocida como análisis funcional. Su aportación al proceso Gram-Schmidt se produjo en un ensayo de 1907 acerca de ecuaciones integrales, en el que escribió los detalles del método de manera más explícita que Gram.

Dicho de manera sucinta, el Teorema 5.15 dice que todo subespacio de $\mathbb{R}^n$ tiene una base ortogonal y proporciona un algoritmo para construir dicha base.

Demostración Se probará por inducción que, para cada $i = 1, \ldots, k$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ es una base ortogonal para W.

Dado que $\mathbf{v}_1 = \mathbf{x}_1$, claramente $\{\mathbf{v}_1\}$ es una base (ortogonal) para $W_1 = \text{gen}(\mathbf{x}_1)$. Ahora suponga que, para algún $i < k$, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ es una base ortogonal para W_i. Entonces

$$\mathbf{v}_{i+1} = \mathbf{x}_{i+1} - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_{i+1}}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{x}_{i+1}}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 - \cdots - \left(\frac{\mathbf{v}_i \cdot \mathbf{x}_{i+1}}{\mathbf{v}_i \cdot \mathbf{v}_i}\right)\mathbf{v}_i$$

Por la hipótesis de inducción, $\{\mathbf{v}_1, \ldots, \mathbf{v}_i\}$ es una base ortogonal para $\text{gen}(\mathbf{x}_1, \ldots, \mathbf{x}_i) = W_i$. En consecuencia,

$$\mathbf{v}_{i+1} = \mathbf{x}_{i+1} - \text{proy}_{W_i}(\mathbf{x}_{i+1}) = \text{perp}_{W_i}(\mathbf{x}_{i+1})$$

De modo que, por el Teorema de Descomposición Ortogonal, $\mathbf{v}_{i+1}$ es ortogonal a W_i. Por definición, $\mathbf{v}_1, \ldots, \mathbf{v}_i$ son combinaciones lineales de $\mathbf{x}_1, \ldots, \mathbf{x}_i$ y, en consecuencia, están en W_i. Por tanto, $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}\}$ es un conjunto ortogonal de vectores en W_{i+1}.

Más aún, $\mathbf{v}_{i+1} \neq \mathbf{0}$, pues de otro modo, $\mathbf{x}_{i+1} = \text{proy}_{W_i}(\mathbf{x}_{i+1})$, que a su vez implica que $\mathbf{x}_{i+1}$ está en W_i. Pero esto es imposible, pues $W_i = \text{gen}(\mathbf{x}_1, \ldots, \mathbf{x}_i)$ y $\{\mathbf{x}_1, \ldots, \mathbf{x}_{i+1}\}$ es linealmente independiente. (¿Por qué?) Se concluye que $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}\}$ es un conjunto de $i + 1$ vectores linealmente independientes en W_{i+1}. En consecuencia, $\{\mathbf{v}_1, \ldots, \mathbf{v}_{i+1}\}$ es una base para W_{i+1}, pues $\dim W_{i+1} = i + 1$. Esto completa la demostración.

Si requiere una base ortonormal para W, simplemente necesita normalizar los vectores ortogonales producidos por el proceso de Gram-Schmidt. Esto es, para cada i, sustituya $\mathbf{v}_i$ por el vector unitario $\mathbf{q}_i = (1/\|\mathbf{v}_i\|)\mathbf{v}_i$.

Ejemplo 5.13

Aplique el proceso de Gram-Schmidt para construir una base ortonormal para el subespacio $W = \text{gen}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3)$ de $\mathbb{R}^4$, donde

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 2 \\ 2 \\ 1 \\ 2 \end{bmatrix}$$

Solución Primero note que $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ es un conjunto linealmente independiente, de modo que forma una base para W. Comience por establecer $\mathbf{v}_1 = \mathbf{x}_1$. A continuación, calcule el componente ortogonal de $\mathbf{x}_2$ a $W_1 = \text{gen}(\mathbf{v}_1)$:

$$\begin{aligned}\mathbf{v}_2 = \text{perp}_{W_1}(\mathbf{x}_2) &= \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_2}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 \\ &= \begin{bmatrix} 2 \\ 1 \\ 0 \\ 1 \end{bmatrix} - \left(\tfrac{2}{4}\right)\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} \frac{3}{2} \\ \frac{3}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}\end{aligned}$$

Para cálculos a mano, es buena idea "escalar" $\mathbf{v}_2$ en este punto para eliminar fracciones. Cuando termine, puede reescalar el conjunto ortogonal que construyó para obtener un conjunto ortonormal; por tanto, puede sustituir cada $\mathbf{v}_i$ por cualquier múltiplo escalar conveniente sin afectar el resultado final. En concordancia, sustituya $\mathbf{v}_2$ por

$$\mathbf{v}_2' = 2\mathbf{v}_2 = \begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix}$$

Ahora encuentre el componente ortogonal de $\mathbf{x}_3$ a

$$W_2 = \text{gen}(\mathbf{x}_1, \mathbf{x}_2) = \text{gen}(\mathbf{v}_1, \mathbf{v}_2) = \text{gen}(\mathbf{v}_1, \mathbf{v}_2')$$

usando la base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2'\}$:

$$\begin{aligned} \mathbf{v}_3 = \text{perp}_{W_2}(\mathbf{x}_3) &= \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2' \cdot \mathbf{x}_3}{\mathbf{v}_2' \cdot \mathbf{v}_2'}\right)\mathbf{v}_2' \\ &= \begin{bmatrix} 2 \\ 2 \\ 1 \\ 2 \end{bmatrix} - \left(\tfrac{1}{4}\right)\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} - \left(\tfrac{15}{20}\right)\begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{2} \\ 0 \\ \frac{1}{2} \\ 1 \end{bmatrix} \end{aligned}$$

De nuevo, reescale y use $\mathbf{v}_3' = 2\mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$.



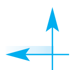
Ahora tiene una base ortogonal $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3'\}$ para W. (Compruebe para asegurarse de que dichos vectores son ortogonales.) Para obtener una base ortonormal, normalice cada vector:

$$\mathbf{q}_1 = \left(\frac{1}{\|\mathbf{v}_1\|}\right)\mathbf{v}_1 = \left(\frac{1}{2}\right)\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$$

$$\mathbf{q}_2 = \left(\frac{1}{\|\mathbf{v}_2'\|}\right)\mathbf{v}_2' = \left(\frac{1}{2\sqrt{5}}\right)\begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3/2\sqrt{5} \\ 3/2\sqrt{5} \\ 1/2\sqrt{5} \\ 1/2\sqrt{5} \end{bmatrix} = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix}$$

$$\mathbf{q}_3 = \left(\frac{1}{\|\mathbf{v}_3'\|}\right)\mathbf{v}_3' = \left(\frac{1}{\sqrt{6}}\right)\begin{bmatrix} -1 \\ 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{6} \\ 0 \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}$$

Entonces $\{\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3\}$ es una base ortonormal para W.



Uno de los usos importantes del proceso de Gram-Schmidt es construir una base ortogonal que contenga un vector especificado. El siguiente ejemplo ilustra esta aplicación.

Ejemplo 5.14 Encuentre una base ortogonal para $\mathbb{R}^3$ que contiene al vector

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Solución Encuentre primero *cualquier* base para $\mathbb{R}^3$ que contenga a $\mathbf{v}_1$. Si se toma

$$\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

entonces $\{\mathbf{v}_1, \mathbf{x}_2, \mathbf{x}_3\}$ es claramente una base para $\mathbb{R}^3$. (¿Por qué?) Ahora aplique el proceso de Gram-Schmidt a esta base para obtener

$$\mathbf{v}_2 = \mathbf{x}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_2}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \left(\tfrac{2}{14}\right)\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{7} \\ \frac{5}{7} \\ -\frac{3}{7} \end{bmatrix}, \quad \mathbf{v}_2' = \begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix}$$

y finalmente

$$\mathbf{v}_3 = \mathbf{x}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{x}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2' \cdot \mathbf{x}_3}{\mathbf{v}_2' \cdot \mathbf{v}_2'}\right)\mathbf{v}_2' = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \left(\tfrac{3}{14}\right)\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \left(\tfrac{-3}{35}\right)\begin{bmatrix} -1 \\ 5 \\ -3 \end{bmatrix} = \begin{bmatrix} \frac{-3}{10} \\ 0 \\ \frac{1}{10} \end{bmatrix},$$

$$\mathbf{v}_3' = \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$$

Entonces $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3'\}$ es una base ortogonal para $\mathbb{R}^3$ que contiene a $\mathbf{v}_1$.

De igual modo, dado un vector unitario, puede encontrar una base ortonormal que lo contenga usando el método anterior y luego normalizar los vectores ortogonales resultantes.

Comentario Cuando se implementa el proceso de Gram-Schmidt en una computadora, casi siempre existe cierto error de redondeo, lo que conduce a una pérdida de ortogonalidad en los vectores $\mathbf{v}_i$. Para evitar esta pérdida de ortogonalidad, usualmente se hacen algunas modificaciones. Los vectores $\mathbf{q}_i$ se normalizan tan pronto como se calculan, en lugar de hacerlo al final, para obtener los vectores $\mathbf{q}_i$ y, conforme se calcula cada $\mathbf{q}_i$, los restantes vectores $\mathbf{x}_j$ se modifican para ser ortogonales a $\mathbf{q}_i$. Este procedimiento se conoce como ***Proceso de Gram-Schmidt modificado***. Sin embargo, en la práctica, se usa una versión de la factorización QR para calcular bases ortonormales.

La factorización *QR*

Si A es una matriz de $m \times n$ con columnas linealmente independientes (lo que requiere que $m \geq n$), entonces, la aplicación del proceso de Gram-Schmidt a dichas columnas produce una factorización de A muy útil que consiste en el producto de una matriz Q con columnas ortonormales y una matriz triangular superior R. Esta es la ***factorización***

QR y tiene aplicaciones en la aproximación numérica de eigenvalores, lo que se explora al final de esta sección, y al problema de aproximación de mínimos cuadrados, que se discute en el capítulo 7.

Para ver cómo surge la factorización QR, sean $\mathbf{a}_1, \ldots, \mathbf{a}_n$ las columnas (linealmente independientes) de A y sean $\mathbf{q}_1, \ldots, \mathbf{q}_n$ los vectores ortonormales obtenidos al aplicar a la matriz A el proceso de Gram-Schmidt con normalizaciones. A partir del Teorema 5.15 se sabe que, para cada $i = 1, \ldots, n$,

$$W_i = \text{gen}(\mathbf{a}_1, \ldots, \mathbf{a}_i) = \text{gen}(\mathbf{q}_1, \ldots, \mathbf{q}_i)$$

Por tanto, existen escalares $r_{1i}, r_{2i}, \ldots, r_{ii}$ tales que

$$\mathbf{a}_i = r_{1i}\mathbf{q}_1 + r_{2i}\mathbf{q}_2 + \cdots + r_{ii}\mathbf{q}_i \text{ para } i = 1, \ldots, n$$

Esto es:

$$\begin{aligned}
\mathbf{a}_1 &= r_{11}\mathbf{q}_1 \\
\mathbf{a}_2 &= r_{12}\mathbf{q}_1 + r_{22}\mathbf{q}_2 \\
&\vdots \\
\mathbf{a}_n &= r_{1n}\mathbf{q}_1 + r_{2n}\mathbf{q}_2 + \cdots + r_{nn}\mathbf{q}_n
\end{aligned}$$

que puede escribirse en forma matricial como

$$A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n] = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_n]\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix} = QR$$

Claramente, la matriz Q tiene columnas ortonormales. También es el caso en que las entradas diagonales de R son todas distintas de cero. Para ver esto, observe que si $r_{ii} = 0$, entonces $\mathbf{a}_i$ es una combinación lineal de $\mathbf{q}_1, \ldots, \mathbf{q}_{i-1}$ y, por tanto, está en W_{i-1}. Pero entonces $\mathbf{a}_i$ sería una combinación lineal de $\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}$, lo que es imposible, pues $\mathbf{a}_1, \ldots, \mathbf{a}_i$ son linealmente independientes. Se concluye que $r_{ii} \neq 0$ para $i = 1, \ldots, n$. Dado que R es triangular superior, debe ser invertible. (Vea el ejercicio 23.)

Acaba de probarse el siguiente teorema.

Teorema 5.16 **La factorización *QR***

Sea A una matriz de $m \times n$ con columnas linealmente independientes. Entonces A puede factorizarse como $A = QR$, donde Q es una matriz de $m \times n$ con columnas ortonormales y R es una matriz triangular superior invertible.

Comentarios

- También puede hacer arreglos para que las entradas diagonales de R sean *positivas*. Si cualquier $r_{ii} < 0$, simplemente sustituya $\mathbf{q}_i$ por $-\mathbf{q}_i$ y r_{ii} por $-r_{ii}$.
- El requisito de que A tenga columnas linealmente independientes es necesario. Para probar esto, suponga que A es una matriz de $m \times n$ que tiene una factorización QR, como en el Teorema 5.16. Entonces, dado que R es invertible, se tiene $Q = AR^{-1}$. Por tanto, rango(Q) = rango(A), por el ejercicio 61 de la sección 3.5. Pero rango$(Q) = n$, pues sus columnas son ortonormales y, en consecuencia, linealmente independientes. De modo que rango$(A) = n$ también, y en consecuencia las columnas de A son linealmente independientes por el teorema fundamental.

- La factorización QR puede extenderse a matrices arbitrarias en una forma ligeramente modificada. Si A es de $m \times n$, se puede encontrar una secuencia de matrices ortogonales $Q_1, \ldots, Q_{m-1}$ tales que $Q_{m-1} \cdots Q_2 Q_1 A$ es una matriz triangular superior R de $m \times n$. Entonces $A = QR$, donde $Q = (Q_{m-1} \cdots Q_2 Q_1)^{-1}$ es una matriz ortogonal. Este planteamiento se examinará en la Exploración: "La factorización QR modificada".

Ejemplo 5.15 Encuentre una factorización QR de

$$A = \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Solución Las columnas de A son justo los vectores del ejemplo 5.13. La base ortonormal para col(A) producida por el proceso de Gram-Schmidt fue

$$\mathbf{q}_1 = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{bmatrix}$$

de modo que

$$Q = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] = \begin{bmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{bmatrix}$$

A partir del Teorema 5.16, $A = QR$ para alguna matriz triangular superior R. Para encontrar R, use el hecho de que Q tiene columnas ortonormales y, en consecuencia, $Q^TQ = I$. Por tanto,

$$Q^TA = Q^TQR = IR = R$$

Calcule

$$R = Q^TA = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ 3\sqrt{5}/10 & 3\sqrt{5}/10 & \sqrt{5}/10 & \sqrt{5}/10 \\ -\sqrt{6}/6 & 0 & \sqrt{6}/6 & \sqrt{6}/3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix}$$

Ejercicios 5.3

En los ejercicios 1-4, los vectores dados forman una base para $\mathbb{R}^2$ o $\mathbb{R}^3$. Aplique el proceso de Gram-Schmidt para obtener una base ortogonal. Luego normalice esta base para obtener una base ortonormal.

1. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

2. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$

3. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 3 \\ 2 \\ 4 \end{bmatrix}$

4. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

En los ejercicios 5 y 6, los vectores dados forman una base para un subespacio W de $\mathbb{R}^3$ *o* $\mathbb{R}^4$*. Aplique el proceso de Gram-Schmidt para obtener una base ortogonal para W.*

5. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$

6. $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \\ 1 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 2 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 1 \\ 8 \\ 1 \\ 0 \end{bmatrix}$

En los ejercicios 7 y 8, encuentre la descomposición ortogonal de **v** *con respecto al subespacio W.*

7. $\mathbf{v} = \begin{bmatrix} 4 \\ -4 \\ 3 \end{bmatrix}$, W como en el ejercicio 5

8. $\mathbf{v} = \begin{bmatrix} 1 \\ 4 \\ 0 \\ 2 \end{bmatrix}$, W como en el ejercicio 6

Use el proceso de Gram-Schmidt para encontrar una base ortogonal para los espacios columna de las matrices en los ejercicios 9 y 10.

9. $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ **10.** $\begin{bmatrix} 1 & 2 & -1 \\ 1 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & 1 & 5 \end{bmatrix}$

11. Encuentre una base ortogonal para $\mathbb{R}^3$ que contenga al vector $\begin{bmatrix} 3 \\ 1 \\ 5 \end{bmatrix}$.

12. Encuentre una base ortogonal para $\mathbb{R}^4$ que contenga a los vectores

$$\begin{bmatrix} 2 \\ 1 \\ 0 \\ -1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 \\ 0 \\ 3 \\ 2 \end{bmatrix}$$

En los ejercicios 13 y 14, encuentre las entradas faltantes de Q para hacer Q una matriz ortogonal.

13. $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & * \\ 0 & 1/\sqrt{3} & * \\ -1/\sqrt{2} & 1/\sqrt{3} & * \end{bmatrix}$

14. $Q = \begin{bmatrix} 1/2 & 2/\sqrt{14} & * & * \\ 1/2 & 1/\sqrt{14} & * & * \\ 1/2 & 0 & * & * \\ 1/2 & -3/\sqrt{14} & * & * \end{bmatrix}$

En los ejercicios 15 y 16, encuentre una factorización QR de la matriz en el ejercicio dado.

15. Ejercicio 9 **16.** Ejercicio 10

En los ejercicios 17 y 18, las columnas de Q se obtuvieron al aplicar el proceso de Gram-Schmidt a las columnas de A. Encuentre la matriz triangular superior R tal que A = QR.

17. $A = \begin{bmatrix} 2 & 8 & 2 \\ 1 & 7 & -1 \\ -2 & -2 & 1 \end{bmatrix}, Q = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}$

18. $A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ -1 & -1 \\ 0 & 1 \end{bmatrix}, Q = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \end{bmatrix}$

19. Si A es una matriz ortogonal, encuentre una factorización QR de A.

20. Demuestre que A es invertible si y sólo si $A = QR$, donde Q es ortogonal y R es triangular superior con entradas distintas de cero en su diagonal.

En los ejercicios 21 y 22, use el método sugerido por el ejercicio 20 para calcular A^{-1} *para la matriz A en el ejercicio dado.*

21. Ejercicio 9 **22.** Ejercicio 15

23. Sea A una matriz de $m \times n$ con columnas linealmente independientes. Proporcione una demostración alternativa de que la matriz triangular superior R en una factorización QR de A debe ser invertible, usando la propiedad (c) del teorema fundamental.

24. Sea A una matriz de $m \times n$ con columnas linealmente independientes y sea $A = QR$ una factorización QR de A. Demuestre que A y Q tienen el mismo espacio columna.

Exploración

La factorización QR modificada

Cuando la matriz A no tiene columnas linealmente independientes, el proceso de Gram-Schmidt no funciona en la forma que se planteó y por tanto no puede usarse para desarrollar una factorización QR generalizada de A. Existe una modificación del proceso de Gram-Schmidt que puede usarse, pero en vez de ello se explorará un método que convierte A en la forma triangular superior columna por columna, usando una secuencia de matrices ortogonales. El método es análogo al de la factorización LU, en el que la matriz L se forma usando una secuencia de matrices elementales.

Lo primero que necesita es el "análogo ortogonal" de una matriz elemental; esto es, se necesita saber cómo construir una matriz ortogonal Q que transformará una columna dada de A, llámela $\mathbf{x}$, en la correspondiente columna de R, llámela $\mathbf{y}$. Por el Teorema 5.6 será necesario que $\|\mathbf{x}\| = \|Q\mathbf{x}\| = \|\mathbf{y}\|$. La figura 5.11 sugiere una forma de proceder: puede reflejar $\mathbf{x}$ en una recta perpendicular a $\mathbf{x} - \mathbf{y}$. Si

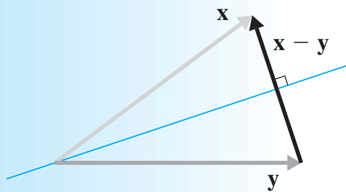


Figura 5.11

$$\mathbf{u} = \left(\frac{1}{\|\mathbf{x} - \mathbf{y}\|}\right)(\mathbf{x} - \mathbf{y}) = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

es el vector unitario en la dirección de $\mathbf{x} - \mathbf{y}$, entonces $\mathbf{u}^{\perp} = \begin{bmatrix} -d_2 \\ d_1 \end{bmatrix}$ es ortogonal a $\mathbf{u}$ y puede usar el ejercicio 26 de la sección 3.6 para encontrar la matriz estándar Q de la reflexión en la recta que pasa por el origen en la dirección de $\mathbf{u}^{\perp}$.

1. Demuestre que $Q = \begin{bmatrix} 1 - 2d_1^2 & -2d_1d_2 \\ -2d_1d_2 & 1 - 2d_2^2 \end{bmatrix} = I - 2\mathbf{u}\mathbf{u}^T$.
2. Calcule Q para

$$\text{(a)}\ \mathbf{u} = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix} \quad \text{(b)}\ \mathbf{x} = \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \mathbf{y} = \begin{bmatrix} 1 \\ 7 \end{bmatrix}$$

Puede generalizar la definición de Q del modo siguiente. Si $\mathbf{u}$ es cualquier vector unitario en $\mathbb{R}^n$, una matriz Q de $n \times n$ se define como

$$Q = I - 2\mathbf{u}\mathbf{u}^T$$

Alston Householder (1904-1993) fue uno de los pioneros en el campo del álgebra lineal numérica. Fue el primero en presentar un tratamiento sistemático de los algoritmos para resolver problemas que involucran sistemas lineales. Además de introducir las ampliamente utilizadas transformaciones de Householder que llevan su nombre, fue uno de los primeros en defender el uso sistemático de normas en el álgebra lineal. Su libro de 1964, *La teoría de matrices en el análisis numérico*, es considerado un clásico.

Tal matriz se llama ***matriz de Householder*** (o un ***reflector elemental***).

3. Demuestre que toda matriz de Householder Q satisface las siguientes propiedades:

(a) Q es simétrica. (b) Q es ortogonal. (c) $Q^2 = I$

4. Pruebe que si Q es una matriz de Householder correspondiente al vector unitario $\mathbf{u}$, entonces

$$Q\mathbf{v} = \begin{cases} -\mathbf{v} & \text{si } \mathbf{v} \text{ está en gen}(\mathbf{u}) \\ \mathbf{v} & \text{si } \mathbf{v} \cdot \mathbf{u} = \mathbf{0} \end{cases}$$

5. Calcule Q para $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ y verifique los problemas 3 y 4.

6. Sea $\mathbf{x} \neq \mathbf{y}$ con $\|\mathbf{x}\| = \|\mathbf{y}\|$ y haga $(1/\|\mathbf{x} - \mathbf{y}\|)(\mathbf{x} - \mathbf{y})$. Demuestre que la correspondiente matriz Q de Householder satisface $Q\mathbf{x} = \mathbf{y}$. [*Sugerencia:* aplique el ejercicio 57 en la sección 1.2 al resultado del problema 4.]

7. Encuentre Q y verifique el problema 6 para

$$\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$$

Ahora está listo para realizar la triangularización de una matriz A de $m \times n$, columna por columna.

8. Sea $\mathbf{x}$ la primera columna de A y sea

$$\mathbf{y} = \begin{bmatrix} \|\mathbf{x}\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Demuestre que si Q_1 es la matriz de Householder dada por el problema 6, entonces Q_1A es una matriz con la forma de bloque

$$Q_1A = \begin{bmatrix} * & * \\ \mathbf{0} & A_1 \end{bmatrix}$$

donde A_1 es $(m - 1)\times (n - 1)$.

Si se repite el problema 8 en la matriz A_1, use una matriz de Householder P_2 tal que

$$P_2A_1 = \begin{bmatrix} * & * \\ \mathbf{0} & A_2 \end{bmatrix}$$

donde A_2 es $(m - 2)\times (n - 2)$.

9. Sea $Q_2 = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & P_2 \end{bmatrix}$. Demuestre que Q_2 es una matriz ortogonal y que

$$Q_2Q_1A = \begin{bmatrix} * & * & * \\ 0 & * & * \\ \mathbf{0} & \mathbf{0} & A_2 \end{bmatrix}$$

10. Demuestre que puede continuar de esta forma para encontrar una sucesión de matrices ortogonales $Q_1, \ldots, Q_{m-1}$ tales que $Q_{m-1} \cdots Q_2 Q_1 A = R$ es una matriz triangular superior de $m \times n$ (es decir, $r_{ij} = 0$ si $i > j$).

11. Deduzca que $A = QR$ con $Q = Q_1 Q_2 \cdots Q_{m-1}$ ortogonal.

12. Use el método de esta exploración para encontrar una factorización QR de

$$\text{(a) } A = \begin{bmatrix} 3 & 9 & 1 \\ -4 & 3 & 2 \end{bmatrix} \quad \text{(b) } A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & -4 & 1 & 1 \\ 2 & -5 & -1 & -2 \end{bmatrix}$$

Cómo aproximar eigenvalores con el algoritmo *QR*

Vea G. H. Golub y C. F. Van Loan, *Matrix Computations* (Baltimore: Johns Hopkins University Press, 1983).

Uno de los mejores métodos (y más ampliamente usados) para aproximar numéricamente los eigenvalores de una matriz, utiliza la factorización QR. El propósito de esta exploración es introducir este método, el ***algoritmo QR***, y demostrarlo en funcionamiento en algunos ejemplos. Para un tratamiento más completo de este tema, consulte cualquier buen libro de álgebra lineal numérica. (Encontrará útil usar un CAS para realizar los cálculos en los problemas siguientes.)

Dada una matriz cuadrada A, el primer paso es factorizarla como $A = QR$ (con cualquier método adecuado). Luego defina $A_1 = RQ$.

1. Demuestre primero que A_1 es semejante a A. Luego demuestre que A_1 tiene los mismos eigenvalores que A.

2. Si $A = \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix}$, encuentre A_1 y verifique que tiene los mismos eigenvalores que A.

Al continuar con el algoritmo, factorice A_1 como $A_1 = Q_1 R_1$ y establezca $A_2 = R_1 Q_1$. Luego factorice $A_2 = Q_2 R_2$ y establezca $A_3 = R_2 Q_2$, etcétera. Esto es: para $k \geq 1$, calcule $A_k = Q_k R_k$ y luego establezca $A_{k+1} = R_k Q_k$.

3. Demuestre que A_k es semejante a A para todo $k \geq 1$.

4. Continuando con el problema 2, calcule A_2, A_3, A_4, y A_5 usando precisión a dos lugares decimales. ¿Qué nota?

Puede demostrarse que si los eigenvalores de A son todos reales y tienen distintos valores absolutos, entonces las matrices A_k tienden a una matriz triangular superior U.

5. ¿Cuáles serán las verdaderas entradas diagonales de esta matriz U?

6. Aproxime los eigenvalores de las siguientes matrices al aplicar el algoritmo QR. Use precisión a dos lugares decimales y realice al menos cinco iteraciones.

$$\text{(a) } \begin{bmatrix} 2 & 3 \\ 2 & 1 \end{bmatrix} \qquad \text{(b) } \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$$

$$\text{(c) } \begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ -4 & 0 & 1 \end{bmatrix} \quad \text{(d) } \begin{bmatrix} 1 & 1 & -1 \\ 0 & 2 & 0 \\ -2 & 4 & 2 \end{bmatrix}$$

7. Aplique el algoritmo QR a la matriz $A = \begin{bmatrix} 2 & 3 \\ -1 & -2 \end{bmatrix}$. ¿Qué ocurre? ¿Por qué?

8. Cambie los eigenvalores de la matriz en el problema 7 al sustituir A con $B = A + 0.9I$. Aplique el algoritmo QR a B y luego cambie de nuevo al restar 0.9 de los eigenvalores (aproximados) de B. Verifique que este método aproxima los eigenvalores de A.

9. Sea $Q_0 = Q$ y $R_0 = R$. Demuestre primero que

$$Q_0 Q_1 \cdots Q_{k-1} A_k = A Q_0 Q_1 \cdots Q_{k-1}$$

para todo $k \geq 1$. Luego demuestre que

$$(Q_0 Q_1 \cdots Q_k)(R_k \cdots R_1 R_0) = A(Q_0 Q_1 \cdots Q_{k-1})(R_{k-1} \cdots R_1 R_0)$$

[*Sugerencia:* use repetidamente el mismo método utilizado para la primera ecuación y trabaje "de adentro hacia afuera".] Finalmente, deduzca que $(Q_0 Q_1 \cdots Q_k)(R_k \cdots R_1 R_0)$ es la factorización QR de A_{k+1}.

5.4 Diagonalización ortogonal de matrices simétricas

En el capítulo 4 se vio que una matriz cuadrada con entradas reales no necesariamente tendrá eigenvalores reales. De hecho, la matriz $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ tiene eigenvalores complejos i y $-i$. También se descubrió que no todas las matrices cuadradas son diagonalizables. La situación cambia dramáticamente si se centra su atención en matrices *simétricas* reales. Como se demostrará en esta sección, todos los eigenvalores de una matriz simétrica real son reales y tal matriz siempre es diagonalizable.

Recuerde que una matriz simétrica es aquella que es igual a su propia transpuesta. Comience por estudiar el proceso de diagonalización para una matriz simétrica de 2×2.

Ejemplo 5.16

Si es posible, diagolanice la matriz $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$.

Solución El polinomio característico de A es $\lambda^2 + \lambda - 6 = (\lambda + 3)(\lambda - 2)$, de donde se ve que A tiene eigenvalores $\lambda_1 = -3$ y $\lambda_2 = 2$. Al calcular los eigenvectores correspondientes se encuentra

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

respectivamente. De modo que A es diagonalizable, y si establece $P = [\mathbf{v}_1 \quad \mathbf{v}_2]$, entonces se sabe que $P^{-1}AP = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix} = D$.

Sin embargo, esto puede mejorarse. Observe que $\mathbf{v}_1$ y $\mathbf{v}_2$ son ortogonales. De modo que si se les normaliza para obtener los eigenvectores unitarios

$$\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix} \quad \text{y} \quad \mathbf{u}_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

y luego se toma

$$Q = [\mathbf{u}_1 \quad \mathbf{u}_2] = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

también se tiene $Q^{-1}AQ = D$. Pero ahora Q es una matriz *ortogonal*, pues $\{\mathbf{u}_1, \mathbf{u}_2\}$ es un conjunto ortonormal de vectores. Por tanto, $Q^{-1} = Q^T$ y se tiene $Q^TAQ = D$. (Note que la comprobación es sencilla, ¡pues calcular Q^{-1} sólo involucra tomar una transpuesta!)

La situación del ejemplo 5.16 es la que tiene interés en este punto. Es lo suficientemente importante para garantizar una nueva definición.

Definición Una matriz cuadrada A es ***diagonalizable ortogonalmente*** si existe una matriz ortogonal Q y una matriz diagonal D tales que $Q^TAQ = D$.

Se está interesado en encontrar condiciones para que una matriz sea diagonalizable ortogonalmente. El Teorema 5.17 muestra dónde buscar.

Teorema 5.17 Si A es diagonalizable ortogonalmente, entonces A es simétrica.

Demostración Si A es diagonalizable ortogonalmente , entonces existe una matriz ortogonal Q y una matriz diagonal D tales que $Q^TAQ = D$. Dado que $Q^{-1} = Q^T$, se tiene $Q^TQ = I = QQ^T$, de modo que

$$QDQ^T = QQ^TAQQ^T = IAI = A$$

Pero entonces

$$A^T = (QDQ^T)^T = (Q^T)^TD^TQ^T = QDQ^T = A$$

pues toda matriz diagonal es simétrica. En consecuencia, A es simétrica.

Comentario El Teorema 5.17 muestra que las matrices diagonalizables ortogonalmente se encuentran en el conjunto de las matrices simétricas. *No* dice que toda matriz simétrica debe ser diagonalizable ortogonalmente. Sin embargo, ¡es notable que esto de hecho sea verdadero! Encontrar una demostración para este resultado sorprendente ocupará gran parte del resto de esta sección.

a + bi

A continuación se probará que no es necesario preocuparse por los eigenvalores *complejos* cuando se trabaja con matrices simétricas con entradas *reales*.

Teorema 5.18 Si A es una matriz simétrica real, entonces los eigenvalores de A son reales.

Recuerde que el *conjugado complejo* de un número complejo $z = a + bi$ es el número $\bar{z} = a - bi$ (vea el Apéndice C). Para demostrar que z es real, es necesario demostrar que $b = 0$. Una forma de hacer esto es demostrar que $z = \bar{z}$, pues entonces $bi = -bi$ (o $2bi = 0$), de donde se tiene que $b = 0$.

La noción de conjugado complejo también se extiende a vectores y matrices, por ejemplo, al definir $\bar{A}$ como la matriz cuyas entradas son los conjugados complejos de las entradas de A; esto es: si $A = [a_{ij}]$, entonces $\bar{A} = [\bar{a}_{ij}]$. Las reglas para la conjugación compleja se extienden fácilmente a matrices; en particular, se tiene $\overline{AB} = \bar{A}\bar{B}$ para matrices compatibles A y B.

Demostración Suponga que λ es un eigenvalor de A con su correspondiente eigenvector $\mathbf{v}$. Entonces $A\mathbf{v} = \lambda\mathbf{v}$ y, al tomar los conjugados complejos, se tiene $\overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}}$. Pero entonces

$$A\bar{\mathbf{v}} = \bar{A}\bar{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$$

pues A es real. Al tomar transpuestas y usar el hecho de que A es simétrica, se tiene

$$\bar{\mathbf{v}}^TA = \bar{\mathbf{v}}^TA^T = (A\bar{\mathbf{v}})^T = (\bar{\lambda}\bar{\mathbf{v}})^T = \bar{\lambda}\bar{\mathbf{v}}^T$$

Por tanto,

$$\lambda(\bar{\mathbf{v}}^T\mathbf{v}) = \bar{\mathbf{v}}^T(\lambda\mathbf{v}) = \bar{\mathbf{v}}^T(A\mathbf{v}) = (\bar{\mathbf{v}}^TA)\mathbf{v} = (\bar{\lambda}\bar{\mathbf{v}}^T)\mathbf{v} = \bar{\lambda}(\bar{\mathbf{v}}^T\mathbf{v})$$

o $(\lambda - \bar{\lambda})(\bar{\mathbf{v}}^T\mathbf{v}) = 0$.

Ahora, si $\mathbf{v} = \begin{bmatrix} a_1 + b_1i \\ \vdots \\ a_n + b_ni \end{bmatrix}$, entonces $\bar{\mathbf{v}} = \begin{bmatrix} a_1 - b_1i \\ \vdots \\ a_n - b_ni \end{bmatrix}$, de modo que

$$\bar{\mathbf{v}}^T\mathbf{v} = (a_1^2 + b_1^2) + \cdots + (a_n^2 + b_n^2) \neq 0$$

pues $\mathbf{v} \neq \mathbf{0}$ (porque es un eigenvector). Se concluye que $\lambda - \bar{\lambda} = 0$, o $\lambda = \bar{\lambda}$. En consecuencia, λ es real.

El Teorema 4.20 demostró que, para cualquier matriz cuadrada, los eigenvectores correspondientes a distintos eigenvalores son linealmente independientes. Para matrices simétricas, algo más fuertes, es verdadero: tales eigenvectores son *ortogonales.*

Teorema 5.19 Si A es una matriz simétrica, entonces cualesquiera dos eigenvectores correspondientes a distintos eigenvalores de A son ortogonales.

Demostración Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ eigenvectores correspondientes a los distintos eigenvalores $\lambda_1 \neq \lambda_2$ de modo que $A\mathbf{v}_1 = \lambda_1\mathbf{v}_1$ y $A\mathbf{v}_2 = \lambda_2\mathbf{v}_2$. Al usar $A^T = A$ y el hecho de que $\mathbf{x}\cdot\mathbf{y} = \mathbf{x}^T\mathbf{y}$ para cualesquiera dos vectores $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$, se tiene

$$\begin{aligned}\lambda_1(\mathbf{v}_1\cdot\mathbf{v}_2) &= (\lambda_1\mathbf{v}_1)\cdot\mathbf{v}_2 = A\mathbf{v}_1\cdot\mathbf{v}_2 = (A\mathbf{v}_1)^T\mathbf{v}_2\\ &= (\mathbf{v}_1^TA^T)\mathbf{v}_2 = (\mathbf{v}_1^TA)\mathbf{v}_2 = \mathbf{v}_1^T(A\mathbf{v}_2)\\ &= \mathbf{v}_1^T(\lambda_2\mathbf{v}_2) = \lambda_2(\mathbf{v}_1^T\mathbf{v}_2) = \lambda_2(\mathbf{v}_1\cdot\mathbf{v}_2)\end{aligned}$$

Por tanto, $(\lambda_1 - \lambda_2)(\mathbf{v}_1\cdot\mathbf{v}_2) = 0$. Pero $\lambda_1 - \lambda_2 \neq 0$, de modo que $\mathbf{v}_1\cdot\mathbf{v}_2 = 0$, como se quería demostrar.

Ejemplo 5.17 Verifique el resultado del Teorema 5.19 para

$$A = \begin{bmatrix} 2 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 2\end{bmatrix}$$

Solución El polinomio característico de A es $-\lambda^3 + 6\lambda^2 - 9\lambda + 4 = -(\lambda - 4)\cdot(\lambda - 1)^2$, de donde se tiene que los eigenvalores de A son $\lambda_1 = 4$ y $\lambda_2 = 1$. Los correspondientes eigenespacios son

$$E_4 = \text{gen}\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right) \quad \text{y} \quad E_1 = \text{gen}\left(\begin{bmatrix}-1\\0\\1\end{bmatrix}, \begin{bmatrix}-1\\1\\0\end{bmatrix}\right)$$


(Compruebe esto.) Fácilmente se verifica que

$$\begin{bmatrix}1\\1\\1\end{bmatrix}\cdot\begin{bmatrix}-1\\0\\1\end{bmatrix} = 0 \quad \text{y} \quad \begin{bmatrix}1\\1\\1\end{bmatrix}\cdot\begin{bmatrix}-1\\1\\0\end{bmatrix} = 0$$


de donde se tiene que todo vector en E_4 es ortogonal a todo vector en E_1. (¿Por qué?)

Comentario Note que $\begin{bmatrix}-1\\0\\1\end{bmatrix}\cdot\begin{bmatrix}-1\\1\\0\end{bmatrix} = 1$. Por tanto, los eigenvectores correspondientes al *mismo* eigenvalor no necesitan ser ortogonales.

Ahora puede demostrar el resultado principal de esta sección. Se le conoce como teorema espectral, pues el conjunto de eigenvalores de una matriz en ocasiones se llaman ***espectro*** de la matriz. (Técnicamente, el Teorema 5.20 debería llamarse teorema espectral *real*, pues existe un resultado correspondiente para matrices con entradas complejas.)

Teorema 5.20 **El teorema espectral**

Sea A una matriz real de $n \times n$. Entonces A es simétrica si y sólo si es diagonalizable ortogonalmente.

Espectro es una palabra latina que significa "imagen". Cuando los átomos vibran, emiten luz. Y cuando la luz pasa a través de un prisma, se dispersa en un espectro: una banda de colores del arco iris. Las frecuencias de vibración corresponden a los eigenvalores de cierto operador y son visibles como líneas brillantes en el espectro de luz que se emite desde un prisma. Por ende, literalmente pueden verse los eigenvalores del átomo en su espectro y, por esta razón, es adecuado que se aplique la palabra *espectro* al conjunto de todos los eigenvalores de una matriz (u operador).

Demostración Ya se demostró la parte "si" como el Teorema 5.17. Para demostrar la implicación "sólo si", se procederá por inducción sobre n. Para $n = 1$, no hay nada que hacer, pues una matriz de 1×1 ya está en forma diagonal. Ahora suponga que toda matriz simétrica real de $k \times k$ con eigenvalores reales es ortogonalmente diagonalizable. Sea $n = k + 1$ y sea A una matriz simétrica real de $n \times n$ con eigenvalores reales.

Sea λ_1 uno de los eigenvalores de A y sea $\mathbf{v}_1$ un eigenvector correspondiente. Entonces $\mathbf{v}_1$ es un vector *real* (¿por qué?) y puede suponer que $\mathbf{v}_1$ es un vector unitario, pues de otro modo puede normalizarse y todavía se tendrá un eigenvector correspondiente a λ_1. Al usar el proceso de Gram-Schmidt, puede extender $\mathbf{v}_1$ a una base ortonormal $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ de $\mathbb{R}^n$. Ahora se forma la matriz

$$Q_1 = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n]$$

Entonces Q_1 es ortogonal y

$$\begin{aligned} Q_1^T A Q_1 &= \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} A[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n] = \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} [A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n] \\ &= \begin{bmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} [\lambda_1\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n] \\ &= \left[\begin{array}{c|c} \lambda_1 & * \\ \hline \mathbf{0} & A_1 \end{array}\right] = B \end{aligned}$$

En una conferencia que impartió en la Universidad de Göttingen en 1905, el matemático alemán David Hilbert (1862-1943) consideró que los operadores lineales actuaban sobre ciertos espacios vectoriales de dimensión infinita. De su conferencia surgió la noción de una forma cuadrática con infinitas variables, y fue en este contexto que Hilbert usó por primera vez el término *espectro* para dar a entender un conjunto completo de eigenvalores. Los espacios en cuestión ahora se llaman *espacios de Hilbert*.

Hilbert realizó grandes aportaciones a muchas áreas de las matemáticas, entre ellas las ecuaciones integrales, la teoría de números, la geometría y los fundamentos de la matemática. En 1900, en el Segundo Congreso Internacional de Matemáticas en París, Hilbert presentó una conferencia titulada "Los problemas de las matemáticas". En ella desafió a los matemáticos a resolver 23 problemas de importancia fundamental durante el siglo venidero. Muchos de los problemas se han resuelto (algunos resultaron verdaderos, otros falsos) y algunos acaso nunca puedan resolverse. No obstante, el discurso de Hilbert alentó a la comunidad matemática y con frecuencia se considera como el discurso más influyente jamás dado acerca de las matemáticas.

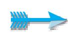
dado que $\mathbf{v}_1^T(\lambda_1\mathbf{v}_1) = \lambda_1(\mathbf{v}_1^T\mathbf{v}_1) = \lambda_1(\mathbf{v}_1 \cdot \mathbf{v}_1) = \lambda_1$ y $\mathbf{v}_i^T(\lambda_1\mathbf{v}_1) = \lambda_1(\mathbf{v}_i^T\mathbf{v}_1) = \lambda_1(\mathbf{v}_i \cdot \mathbf{v}_1) = 0$ para $i \neq 1$, porque $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es un conjunto ortonormal.

Pero

$$B^T = (Q_1^TAQ_1)^T = Q_1^TA^T(Q_1^T)^T = Q_1^TAQ_1 = B$$

de modo que B es simétrica. En consecuencia, B tiene la forma de bloque

$$B = \left[\begin{array}{c|c} \lambda_1 & \mathbf{0} \\ \hline \mathbf{0} & A_1 \end{array}\right]$$


y A_1 es simétrica. Más aún, B es semejante a A (¿por qué?), de modo que el polinomio característico de B es igual al polinomio característico de A, por el Teorema 4.22. Por el ejercicio 39 de la sección 4.3, el polinomio característico de A_1 divide al polinomio característico de A. Se tiene que los eigenvalores de A_1 también son eigenvalores de A y, en consecuencia, son reales. También se ve que A_1 tiene entradas reales. (¿Por qué?) Por

tanto, A_1 es una matriz simétrica real de $k \times k$ con eigenvalores reales, de modo que puede aplicar la hipótesis de inducción. Por tanto, existe una matriz ortogonal P_2 tal que $P_2^TA_1P_2$ es una matriz diagonal; por decir, D_1. Ahora sea

$$Q_2 = \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & P_2 \end{array}\right]$$

Entonces Q_2 es una matriz ortogonal $(k + 1)\times(k + 1)$ y por tanto también lo es $Q = Q_1Q_2$. En consecuencia,

$$\begin{aligned} Q^TAQ &= (Q_1Q_2)^TA(Q_1Q_2) = (Q_2^TQ_1^T)A(Q_1Q_2) = Q_2^T(Q_1^TAQ_1)Q_2 = Q_2^TBQ_2 \\ &= \left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & P_2^T \end{array}\right]\left[\begin{array}{c|c} \lambda_1 & \mathbf{0} \\ \hline \mathbf{0} & A_1 \end{array}\right]\left[\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & P_2 \end{array}\right] \\ &= \left[\begin{array}{c|c} \lambda_1 & \mathbf{0} \\ \hline \mathbf{0} & P_2^TA_1P_2 \end{array}\right] \\ &= \left[\begin{array}{c|c} \lambda_1 & \mathbf{0} \\ \hline \mathbf{0} & D_1 \end{array}\right] \end{aligned}$$

que es una matriz diagonal. Esto completa el paso de inducción y se concluye que, para todo $n \geq 1$, una matriz simétrica real de $n \times n$ con eigenvalores reales es diagonalizable ortogonalmente.

Ejemplo 5.18

Diagonalice ortogonalmente la matriz

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

Solución Ésta es la matriz del ejemplo 5.17. Ya se encontró que los eigenespacios de A son

$$E_4 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) \quad \text{y} \quad E_1 = \text{gen}\left(\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\right)$$

Se necesitan tres eigenvectores ortonormales. Primero, aplique el proceso de Gram-Schmidt a

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

para obtener

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{bmatrix}$$

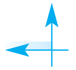


El nuevo vector, que se construyó para ser ortogonal a $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$, todavía está en E_1 (¿por qué?) y es ortogonal a $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Por ende, se tienen tres vectores mutuamente ortogonales, y todo lo que necesita hacer es normalizarlos y construir una matriz Q con estos vectores como sus columnas. Se descubre que

$$Q = \begin{bmatrix} 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix}$$

y es fácil verificar que

$$Q^TAQ = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



El teorema espectral permite escribir una matriz simétrica real A en la forma $A = QDQ^T$, donde Q es ortogonal y D es diagonal. Las entradas diagonales de D son justo los eigenvalores de A, y si las columnas de Q son los vectores ortonormales $\mathbf{q}_1, \ldots, \mathbf{q}_n$, entonces, al usar la representación columna-renglón del producto, se tiene

$$\begin{aligned} A = QDQ^T &= [\mathbf{q}_1 \ \cdots \ \mathbf{q}_n] \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \\ &= [\lambda_1\mathbf{q}_1 \ \cdots \ \lambda_n\mathbf{q}_n] \begin{bmatrix} \mathbf{q}_1^T \\ \vdots \\ \mathbf{q}_n^T \end{bmatrix} \\ &= \lambda_1\mathbf{q}_1\mathbf{q}_1^T + \lambda_2\mathbf{q}_2\mathbf{q}_2^T + \cdots + \lambda_n\mathbf{q}_n\mathbf{q}_n^T \end{aligned}$$

A esto se le conoce como la ***descomposición espectral*** de A. Cada uno de los términos $\lambda_i\mathbf{q}_i\mathbf{q}_i^T$ es una matriz de rank 1, por el ejercicio 62 de la sección 3.5, y $\mathbf{q}_i\mathbf{q}_i^T$ es en realidad la matriz de la proyección sobre el subespacio generado por $\mathbf{q}_i$. (Vea el ejercicio 25.) Por esta razón, la descomposición espectral

$$A = \lambda_1\mathbf{q}_1\mathbf{q}_1^T + \lambda_2\mathbf{q}_2\mathbf{q}_2^T + \cdots + \lambda_n\mathbf{q}_n\mathbf{q}_n^T$$

en ocasiones se conoce como la ***forma en proyección del teorema espectral***.

Ejemplo 5.19

Encuentre la descomposición espectral de la matriz A del ejemplo 5.18.

Solución Del ejemplo 5.18, se tiene:

$$\lambda_1 = 4, \qquad \lambda_2 = 1, \qquad \lambda_3 = 1$$

$$\mathbf{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix}$$

Por tanto,

$$\mathbf{q}_1\mathbf{q}_1^T = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} = \begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

$$\mathbf{q}_2\mathbf{q}_2^T = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & -1/2 \\ 0 & 0 & 0 \\ -1/2 & 0 & 1/2 \end{bmatrix}$$

$$\mathbf{q}_3\mathbf{q}_3^T = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ -1/\sqrt{6} \end{bmatrix} \begin{bmatrix} -1/\sqrt{6} & 2/\sqrt{6} & -1/\sqrt{6} \end{bmatrix} = \begin{bmatrix} 1/6 & -1/3 & 1/6 \\ -1/3 & 2/3 & -1/3 \\ 1/6 & -1/3 & 1/6 \end{bmatrix}$$

de modo que

$$\begin{aligned} A &= \lambda_1\mathbf{q}_1\mathbf{q}_1^T + \lambda_2\mathbf{q}_2\mathbf{q}_2^T + \lambda_3\mathbf{q}_3\mathbf{q}_3^T \\ &= 4\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} + \begin{bmatrix} \frac{1}{2} & 0 & -\frac{1}{2} \\ 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} + \begin{bmatrix} \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ \frac{1}{6} & -\frac{1}{3} & \frac{1}{6} \end{bmatrix} \end{aligned}$$

que puede verificarse con facilidad.

En este ejemplo, $\lambda_2 = \lambda_3$, así que se combinarían los últimos dos términos $\lambda_2\mathbf{q}_2\mathbf{q}_2^T + \lambda_3\mathbf{q}_3\mathbf{q}_3^T$ para obtener

$$\begin{bmatrix} \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{1}{3} & -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

La matriz $\mathbf{q}_2\mathbf{q}_2^T + \mathbf{q}_3\mathbf{q}_3^T$ de rank 2 es la matriz de una proyección sobre el subespacio bidimensional (es decir, el plano) generado por $\mathbf{q}_2$ y $\mathbf{q}_3$. (Vea el ejercicio 26.)

Observe que la descomposición espectral expresa una matriz simétrica A de manera explícita en términos de sus eigenvalores y eigenvectores. Esto proporciona una forma de construir una matriz con eigenvalores dados y eigenvectores (ortonormales).

Ejemplo 5.20

Encuentre una matriz de 2×2 con eigenvalores $\lambda_1 = 3$ y $\lambda_2 = -2$ y correspondientes eigenvectores

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 4 \end{bmatrix} \quad \text{y} \quad \mathbf{v}_2 = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$$

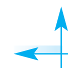
Solución Comience por normalizar los vectores para obtener una base ortonormal $\{\mathbf{q}_1, \mathbf{q}_2\}$,con

$$\mathbf{q}_1 = \begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix} \quad \text{y} \quad \mathbf{q}_2 = \begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}$$

Ahora, calcule la matriz A cuya descomposición espectral es

$$\begin{aligned} A &= \lambda_1\mathbf{q}_1\mathbf{q}_1^T + \lambda_2\mathbf{q}_2\mathbf{q}_2^T \\ &= 3\begin{bmatrix} \frac{3}{5} \\ \frac{4}{5} \end{bmatrix}\begin{bmatrix} \frac{3}{5} & \frac{4}{5} \end{bmatrix} - 2\begin{bmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{bmatrix}\begin{bmatrix} -\frac{4}{5} & \frac{3}{5} \end{bmatrix} \\ &= 3\begin{bmatrix} \frac{9}{25} & \frac{12}{25} \\ \frac{12}{25} & \frac{16}{25} \end{bmatrix} - 2\begin{bmatrix} \frac{16}{25} & -\frac{12}{25} \\ -\frac{12}{25} & \frac{9}{25} \end{bmatrix} \\ &= \begin{bmatrix} -\frac{1}{5} & \frac{12}{5} \\ \frac{12}{5} & \frac{6}{5} \end{bmatrix} \end{aligned}$$


Es fácil comprobar que A tiene las propiedades deseadas. (Hágalo.)



Ejercicios 5.4

Diagonalice ortogonalmente las matrices de los ejercicios 1-10 al encontrar una matriz ortogonal Q y una matriz diagonal D tales que $Q^TAQ = D$.

1. $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & -2 \\ -2 & 1 \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{bmatrix}$

4. $A = \begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}$

5. $A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & 1 \end{bmatrix}$

6. $A = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 4 \\ 2 & 4 & 2 \end{bmatrix}$

7. $A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

8. $A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$

9. $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

10. $A = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 3 \end{bmatrix}$

11. Si $b \neq 0$, diagonalice ortogonalmente $A = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$.

12. Si $b \neq 0$, diagonalice ortogonalmente $A = \begin{bmatrix} a & 0 & b \\ 0 & a & 0 \\ b & 0 & a \end{bmatrix}$.

13. Sean A y B matrices de $n \times n$ diagonalizables ortogonalmente y sea c un escalar. Use el teorema espectral para probar que las siguientes matrices son diagonalizables ortogonalmente:

(a) $A + B$ **(b)** cA **(c)** A^2

14. Si A es una matriz invertible que es diagonalizable ortogonalmente, demuestre que A^{-1} es diagonalizable ortogonalmente.

15. Si A y B son diagonalizables ortogonalmente y $AB = BA$, demuestre que AB es diagonalizable ortogonalmente.

16. Si A es una matriz simétrica, demuestre que todo eigenvalor de A es no negativo si y sólo si $A = B^2$ para alguna matriz simétrica B.

En los ejercicios 17-20, encuentre una descomposición espectral de la matriz en el ejercicio dado.

17. Ejercicio 1 **18.** Ejercicio 2

19. Ejercicio 5 **20.** Ejercicio 8

En los ejercicios 21 y 22, encuentre una matriz de 2 × 2 con eigenvalores λ_1 y λ_2 y correspondientes eigenvectores ortogonales $\mathbf{v}_1$ *y* $\mathbf{v}_2$.

21. $\lambda_1 = -1, \lambda_2 = 2, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

22. $\lambda_1 = 2, \lambda_2 = -2, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$

En los ejercicios 23 y 24, encuentre una matriz simétrica de 3 × 3 con eigenvalores λ_1, λ_2 *y* λ_3 *y correspondientes eigenvectores ortogonales* $\mathbf{v}_1$, $\mathbf{v}_2$ *y* $\mathbf{v}_3$.

23. $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix},$

$\mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}$

24. $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 2, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix},$

$\mathbf{v}_3 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$

25. Sea $\mathbf{q}$ un vector unitario en $\mathbb{R}^n$ y sea W el subespacio generado por $\mathbf{q}$. Demuestre que la proyección ortogonal de un vector $\mathbf{v}$ sobre W (como se definió en las secciones 1.2 y 5.2) está dada por

$$\text{proy}_W(\mathbf{v}) = (\mathbf{q}\mathbf{q}^T)\mathbf{v}$$

y que la matriz de esta proyección es, por tanto, $\mathbf{q}\mathbf{q}^T$. [*Sugerencia:* recuerde que para $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$, $\mathbf{x} \cdot \mathbf{y} = \mathbf{x}^T\mathbf{y}$.]

26. Sea $\{\mathbf{q}_1, \ldots, \mathbf{q}_k\}$ un conjunto ortonormal de vectores en $\mathbb{R}^n$ y sea W el subespacio generado por este conjunto.

(a) Demuestre que la matriz de la proyección ortogonal sobre W está dada por

$$P = \mathbf{q}_1\mathbf{q}_1^T + \cdots + \mathbf{q}_k\mathbf{q}_k^T$$

(b) Demuestre que la matriz de proyección P en el inciso (a) es simétrica y satisface $P^2 = P$.

(c) Sea $Q = [\mathbf{q}_1 \ \cdots \ \mathbf{q}_k]$ la matriz de $n \times k$ cuyas columnas son los vectores base ortonormales de W. Demuestre que $P = QQ^T$ y deduzca que rank(P) $= k$.

27. Sea A una matriz real de $n \times n$, cuyos eigenvalores son todos reales. Demuestre que existe una matriz ortogonal Q y una matriz triangular superior T tales que $Q^TAQ = T$. Este muy útil resultado se conoce como ***teorema de Triangularización de Schur***. [*Sugerencia:* adapte la prueba del teorema espectral.]

28. Sea A una matriz nilpotente (vea el ejercicio 56 de la sección 4.2). Demuestre que existe una matriz ortogonal Q tal que $Q^T AQ$ es triangular superior con ceros en su diagonal. [*Sugerencia:* use el ejercicio 27.]

5.5 Aplicaciones

Códigos duales

Existen muchas formas de construir nuevos códigos a partir de los antiguos. En esta sección se considerará uno de los más importantes de ellos.

Primero, es necesario generalizar los conceptos de un generador y una matriz de control de paridad para un código. Recuerde de la sección 3.7 que una matriz generadora estándar para un código es una matriz de $n \times m$ de la forma

$$G = \begin{bmatrix} I_m \\ \hdashline A \end{bmatrix}$$

y una matriz de verificación de paridad estándar es una matriz $(n - m) \times n$ de la forma

$$P = [B \mid I_{n-m}]$$



Observe que la forma de dichas matrices garantiza que las columnas de G y los renglones de P sean linealmente independientes. (¿Por qué?) Al demostrar el Teorema 3.37 se

demostró que G y P están asociadas con el *mismo* código si y sólo si $A = B$, lo que equivale a requerir que $PG = O$. Use estas propiedades como la base para la siguiente definición.

Definición Para $n > m$, una matriz G de $n \times m$ y una matriz P de $(n - k) \times n$ (con entradas en $\mathbb{Z}_2$) son una ***matriz generadora*** y una ***matriz de control de paridad***, respectivamente, para un código C binario (n, k) si se satisfacen todas las condiciones siguientes:

1. Las columnas de G son linealmente independientes.
2. Los renglones de P son linealmente independientes.
3. $PG = O$

Note que la propiedad (3) implica que toda columna $\mathbf{v}$ de G satisface $P\mathbf{v} = \mathbf{0}$ y por tanto es un vector código en C. Además, un vector $\mathbf{y}$ está en C si y sólo si se obtiene a partir de la matriz generadora como $\mathbf{y} = G\mathbf{u}$ para algún vector $\mathbf{u}$ en $\mathbb{Z}_2^n$. En otras palabras, C es el espacio columna de G.

Para entender la relación entre diferentes matrices generadoras para el mismo código, sólo necesita recordar que así como las operaciones elementales con renglón no afectan el espacio renglón de una matriz (por el Teorema 3.20), las operaciones elementales en columna no afectan el espacio columna. Para una matriz sobre $\mathbb{Z}_2$, sólo existen dos operaciones relevantes: intercambio de dos columnas (C1) y la adición de una columna a otra (C2).  (¿Por qué estas son las únicas dos operaciones elementales en columna para matrices en $\mathbb{Z}_2$?)

De igual modo, las operaciones elementales con renglón conservan la independencia lineal de los renglones de P. Más aún: si E es una matriz elemental y $\mathbf{c}$ es un vector código, entonces

$$(EP)\mathbf{c} = E(P\mathbf{c}) = E\mathbf{0} = \mathbf{0}$$

Se tiene que EP también es una matriz de control de paridad para C. En consecuencia, cualquier matriz de control de paridad puede convertirse en otra mediante una secuencia de operaciones con renglón: intercambiar dos renglones (R1) y sumar un renglón a otro (R2). Se está interesado en demostrar que cualquier matriz generadora o de control de paridad puede llevarse a forma estándar. Existe otra definición que se necesita. Dos códigos C_1 y C_2 se llamarán ***equivalentes*** si existe una matriz permutación M tal que

$$\{M\mathbf{x} : \mathbf{x} \text{ en } C_1\} = C_2$$

En otras palabras, si se permutan las entradas de los vectores en C_1 (todas en la misma forma), puede obtenerse C_2. Por ejemplo,

$$C_1 = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \quad \text{y} \quad C_2 = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

son equivalentes vía la matriz permutación $M = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. Permutar las entradas de vectores código corresponde a permutar los *renglones* de una matriz generadora y permutar las *columnas* una matriz de control de paridad para el código.  (¿Por qué?)

Es posible llevar una matriz generadora para un código a forma estándar mediante las operaciones C1, C2 y R1. Si no se usó R1, entonces se tiene el mismo código; si se usó R1, entonces se tiene un código equivalente. Puede llevar cualquier matriz de control de paridad para un código a forma estándar mediante las operaciones R1, R2 y C1. Si no se usó C1, entonces se tiene el mismo código; si se usó C1, entonces se tiene un código equivalente.

Los siguientes ejemplos ilustran estos puntos.

Ejemplo 5.21 (a) Lleve la matriz generadora

$$G = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

a la forma estándar y encuentre una matriz de control de paridad asociada.

(b) Lleve la matriz de control de paridad

$$P = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

a la forma estándar y encuentre una matriz generadora asociada.

Solución (a) Puede llevar la matriz generadora G a la forma estándar del modo siguiente:

$$G = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} I \\ \hdashline A \end{bmatrix} = G'$$

(¿Ve por qué no es posible obtener una forma estándar sin usar R1?) Por tanto, $A = [1 \quad 0]$, de modo que

$$P = [A \,\vdots\, I] = [1 \quad 0 \quad 1]$$

es una matriz de control de paridad asociada, por el Teorema 3.37.

(b) Use operaciones elementales por renglones para llevar P a la forma estándar y tenga en mente que se quiere crear una matriz identidad a la *derecha*, no a la izquierda, como en la eliminación de Gauss-Jordan. Calcule

$$P = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \xrightarrow{R_1 + R_2} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} = [A \,\vdots\, I] = P'$$

Por tanto, $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, de modo que, por el Teorema 3.37, una matriz generadora asociada es

$$G = \begin{bmatrix} I \\ \hdashline A \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$$

Comentario En el inciso (a), es instructivo verificar que G y G' generan códigos equivalentes, mas no idénticos. Compruebe que esto es así al calcular $\{G\mathbf{x} : \mathbf{x} \text{ en } \mathbb{Z}_2^2\}$ y $\{G'\mathbf{x} : \mathbf{x} \text{ en } \mathbb{Z}_2^2\}$.

Ahora regrese su atención al tema principal de esta sección: la noción de código dual.

Definición Sea C un conjunto de vectores código en $\mathbb{Z}_2^n$. El complemento ortogonal de C se llama ***código dual*** de C y se denota $C^\perp$. Esto es,

$$C^\perp = \{\mathbf{x} \text{ en } \mathbb{Z}_2^n : \mathbf{c} \cdot \mathbf{x} = 0 \quad \text{para todo } \mathbf{c} \text{ en } C\}$$

El producto punto en $\mathbb{Z}_2^n$ se comporta como el producto punto en $\mathbb{R}^n$, con una excepción importante: la propiedad (d) del Teorema 1.2 ya no es verdadero. En otras palabras, en $\mathbb{Z}_2^n$, ¡un vector distinto de cero puede ser ortogonal a sí mismo! Como un ejemplo, considere $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ en $\mathbb{Z}_2^2$. Entonces $\mathbf{x} \cdot \mathbf{x} = 1 + 1 = 0$.

Ejemplo 5.22 Encuentre el código dual del código C en el ejemplo 5.21(b).

Solución El código C es

$$C = \{G\mathbf{x} : \mathbf{x} \text{ en } \mathbb{Z}_2^2\}$$

$$= \left\{ G\begin{bmatrix} 0 \\ 0 \end{bmatrix}, G\begin{bmatrix} 0 \\ 1 \end{bmatrix}, G\begin{bmatrix} 1 \\ 0 \end{bmatrix}, G\begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$



[De manera alternativa, $C = \{\mathbf{c} \text{ en } \mathbb{Z}^4 : P\mathbf{c} = \mathbf{0}\} = \text{nulo}(P)$. Compruebe que esto realmente produce el mismo código.]

Para encontrar $C^\perp$, es necesario que dichos vectores en $\mathbb{Z}_2^4$ que son ortogonales a los otros cuatro vectores en C. Puesto que sólo existen 16 vectores en $\mathbb{Z}_2^4$, podría proceder por ensayo y error, pero aquí hay un método mejor. Sea $\mathbf{y} = [y_1 \quad y_2 \quad y_3 \quad y_4]^T$ que está en $C^\perp$. Puesto que $\mathbf{y} \cdot \mathbf{c} = 0$ para cada $\mathbf{c}$ en C, se tienen cuatro ecuaciones, una de las cuales puede ignorarse, pues sólo dice que $0 = 0$. Las otras tres son

$$\begin{aligned} y_2 + y_3 &= 0 \\ y_1 + y_3 + y_4 &= 0 \\ y_1 + y_2 + y_4 &= 0 \end{aligned}$$

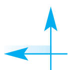
Al resolver este sistema se obtiene

$$\left[\begin{array}{cccc|c} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|c} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$



(Compruebe esto.) Se tiene que $y_1 = y_3 + y_4$ y $y_2 = y_3$, de modo que

$$C^\perp = \left\{ \begin{bmatrix} y_3 + y_4 \\ y_3 \\ y_3 \\ y_4 \end{bmatrix} \right\} = \left\{ y_3 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + y_4 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$



Ahora se examina la relación entre las matrices generadora y de control de paridad de un código y su dual.

Teorema 5.21 Si C es un código binario (n, k) con matriz generadora G y matriz de control de paridad P, entonces $C^\perp$ es un código binario $(n, n-k)$ tal que

a. G^T es una matriz de verificación de paridad para $C^\perp$.
b. P^T es una matriz generadora para $C^\perp$.

Demostración Por definición, G es una matriz de $n \times k$ con columnas linealmente independientes, P es una matriz $(n-k) \times n$ con renglones linealmente independientes y $PG = O$. Por tanto, los renglones de G^T son linealmente independientes, las columnas de P^T son linealmente independientes y

$$G^TP^T = (PG)^T = O^T = O$$

Esto demuestra que G^T es una matriz de control de paridad para $C^\perp$ y P^T es una matriz generadora para $C^\perp$. Dado que P^T es $n \times (n-k)$, $C^\perp$ es un código $(n, n-k)$.

Ejemplo 5.23 Encuentre las matrices generadora y de verificación de paridad para el código dual $C^\perp$ del ejemplo 5.22.

Solución Existen dos formas de proceder. Se ilustrarán ambos métodos.

Método 1: de acuerdo con el Teorema 5.21(b), una matriz generadora $G^\perp$ para $C^\perp$ está dada por

$$G^\perp = P^T = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Esta matriz está en forma estándar con $A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$, de modo que una matriz de control de paridad para $C^\perp$ es

$$P^\perp = [A \mid I] = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

por el Teorema 3.37.

Método 2: con el Teorema 5.21(a) y con referencia al ejemplo 5.21(b), se obtiene una matriz de control de paridad $P^\perp$ para $C^\perp$ del modo siguiente:

$$P^\perp = G^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

Esta matriz no está en forma estándar, así que se usan operaciones elementales con renglones para convertirla en

$$P^\perp = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix} = [A \mid I] = P_1^\perp$$

Ahora puede usar el Teorema 3.37 para obtener una matriz generadora $G^\perp$ para $C^\perp$:

$$G^\perp = \begin{bmatrix} I \\ \hdashline A \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

Ejemplo 5.24

Sea C el código con matriz generadora

$$G = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Mencione los vectores en C y $C^\perp$.

Solución El código C es

$$C = \{G\mathbf{x} : \mathbf{x} \text{ en } \mathbb{Z}_2^4\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

(Note que C es un código de doble repetición que codifica vectores de $\mathbb{Z}_2^2$ como vectores en $\mathbb{Z}_2^4$ al escribir las entradas dos veces. Vea el ejercicio 79 de la sección 3.7.) Con el Teorema 5.21, se encuentra que la matriz de control de paridad $P^\perp$ para $C^\perp$ es

$$P^\perp = G^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}^T = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

Por ende, $P^\perp$ tiene la forma $[A \mid I]$, donde $A = I$, de modo que una matriz generadora $G^\perp$ para $C^\perp$ es

$$G^\perp = \begin{bmatrix} I \\ \hdashline A \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = G$$

En consecuencia, $C^\perp$ tiene la misma matriz generadora que C, ¡de modo que $C^\perp = C$!

Un código C con la propiedad de que $C^\perp = C$ se llama ***autodual***. Puede comprobar que el código del ejemplo 5.24 es autodual al demostrar que todo vector en C es ortogonal a todos los vectores en C, incluido él mismo. (Hágalo.)

Acaso haya notado que, en el código autodual del ejemplo 5.24 todo vector en C tiene un número *par* de números 1. Se demostrará que esto es verdadero para todo código autodual. La siguiente definición es útil.

F. Jessie MacWilliams (1917-1990) fue una de las pioneras en la teoría de codificación. Recibió su licenciatura y maestría en la Cambridge University en 1938-39, después de lo cual estudió en Estados Unidos, en la Johns Hopkins University y en Harvard University. Después de casarse y criar a su familia, MacWilliams tomó un puesto como programadora de computadoras en los Bell Laboratories en Murray Hill, New Jersey, en 1958, donde comenzó a interesarse en la teoría de codificación. En 1961 regresó a Harvard durante un año y obtuvo su doctorado.

Su tesis contiene uno de los teoremas más poderosos de la teoría de codificación. Ahora conocido como las **identidades MacWilliams**, este teorema relaciona la distribución ponderada (el número de palabras código de cada peso posible) de un código lineal, con la distribución ponderada de su código dual. Los teóricos en codificación usan ampliamente las identidades MacWilliams tanto para obtener nueva información teórica acerca de los códigos de corrección de errores, como para determinar las distribuciones ponderadas de códigos específicos.

MacWilliams acaso es mejor conocida por su libro *The Theory of Error-Correcting Codes* (1977), escrito con N. J. A. Sloane de Bell Labs. Con frecuencia este libro se conoce como "la biblia de la teoría de codificación". En 1980, MacWilliams pronunció el discurso inaugural Emmy Noether de la Asociación de Mujeres en Matemáticas.

Definición Sea $\mathbf{x}$ un vector en $\mathbb{Z}_2^n$. El ***peso*** de $\mathbf{x}$, denotado $w(\mathbf{x})$, es el número de números 1 en $\mathbf{x}$.

Por ejemplo, $w([1 \quad 1 \quad 0 \quad 1 \quad 0 \quad 0 \quad 1]^T) = 4$. Si temporalmente considera $\mathbf{x}$ como un vector en $\mathbb{R}^n$, entonces pueden proporcionarse las siguientes descripciones alternativas de $w(\mathbf{x})$. Sea $\mathbf{1}$ que denota al vector (de la misma longitud que $\mathbf{x}$) cuyas entradas son todas 1. Entonces $w(\mathbf{x}) = \mathbf{x} \cdot \mathbf{1}$ y $w(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x}$. Ahora puede probar los siguientes hechos interesantes acerca de los códigos autoduales.

Teorema 5.22 Si C es un código autodual, entonces:

a. Todo vector en C tiene peso par.
b. $\mathbf{1}$ está en C.

Demostración (a) Un vector $\mathbf{x}$ en $\mathbb{Z}_2^n$ tiene peso par si y sólo si $w(\mathbf{x}) = 0$ en $\mathbb{Z}_2$. Pero

$$w(\mathbf{x}) = \mathbf{x} \cdot \mathbf{x} = 0$$

pues C es autodual.

(b) Con la propiedad (a), se tiene $\mathbf{1} \cdot \mathbf{x} = w(\mathbf{x}) = 0$ en $\mathbb{Z}_2$ para toda $\mathbf{x}$ en C. Esto significa que $\mathbf{1}$ es ortogonal a todo vector en C, de modo que $\mathbf{1}$ está en $C^{\perp} = C$, como se requiere.

Formas cuadráticas

Una expresión de la forma

$$ax^2 + by^2 + cxy$$

se llama ***forma cuadrática*** en x y y. De igual modo,

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz$$

es una forma cuadrática en x, y y z. En palabras: una forma cuadrática es una suma de términos, cada uno de los cuales tiene grado *dos*. Por tanto, $5x^2 - 3y^2 + 2xy$ es una forma cuadrática, pero $x^2 + y^2 + x$ no lo es.

Las formas cuadráticas pueden representarse usando matrices del modo siguiente:

$$ax^2 + by^2 + cxy = [x \quad y]\begin{bmatrix} a & c/2 \\ c/2 & b \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}$$

y

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz = [x \quad y \quad z]\begin{bmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

(Verifique esto.) Cada una tiene la forma $\mathbf{x}^T A\mathbf{x}$, donde la matriz A es simétrica. Esta observación conduce a la siguiente definición general.

Definición Una ***forma cuadrática*** con n variables es una función $f: \mathbb{R}^n \to \mathbb{R}$ de la forma

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$$

donde A es una matriz simétrica de $n \times n$ y $\mathbf{x}$ está en $\mathbb{R}^n$. A se conoce como ***matriz asociada con f***.

Ejemplo 5.25

¿Cuál es la forma cuadrática con matriz asociada $A = \begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}$?

Solución Si $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, entonces

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} = [x_1 \quad x_2]\begin{bmatrix} 2 & -3 \\ -3 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 5x_2^2 - 6x_1x_2$$

Observe que las entradas *fuera de la diagonal* $a_{12} = a_{21} = -3$ de A se *combinan* para producir el coeficiente -6 de x_1x_2. En general esto es verdadero. Es posible desarrollar una forma cuadrática con n variables $\mathbf{x}^T A\mathbf{x}$ del modo siguiente:

$$\mathbf{x}^T A\mathbf{x} = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + \sum_{i<j} 2a_{ij}x_ix_j$$

Por tanto, si $i \neq j$, el coeficiente de x_ix_j es $2a_{ij}$.

Ejemplo 5.26

Encuentre la matriz asociada con la forma cuadrática

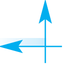
$$f(x_1, x_2, x_3) = 2x_1^2 - x_2^2 + 5x_3^2 + 6x_1x_2 - 3x_1x_3$$

Solución Los coeficientes de los términos al cuadrado x_i^2 pasan a la diagonal como a_{ii}, y los coeficientes de los términos en x_ix_j se dividen entre a_{ij} y a_{ji}. Esto produce

$$A = \begin{bmatrix} 2 & 3 & -\frac{3}{2} \\ 3 & -1 & 0 \\ -\frac{3}{2} & 0 & 5 \end{bmatrix}$$

de este modo $$f(x_1, x_2, x_3) = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 2 & 3 & -\frac{3}{2} \\ 3 & -1 & 0 \\ -\frac{3}{2} & 0 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

como puede comprobar fácilmente.



En el caso de una forma cuadrática $f(x, y)$ con dos variables, la gráfica de $z = f(x, y)$ es una superficie en $\mathbb{R}^3$. En la figura 5.12 se muestran algunos ejemplos.

Observe que el efecto de mantener x o y constante es tomar una sección transversal de la gráfica paralela a los planos yz o xz, respectivamente. Para las gráficas de la figura 5.12, todas estas secciones transversales son fáciles de identificar. Por ejemplo, en la figura 5.12(a), las secciones transversales que se obtienen al mantener x o y constante son todas parábolas que se abren hacia arriba, de modo que $f(x, y) \geq 0$ para todos los valores de x y y. En la figura 5.12(c), mantener x constante produce parábolas que se abren hacia abajo y mantener y constante produce parábolas que se abren hacia arriba, lo que produce un *punto de silla*.

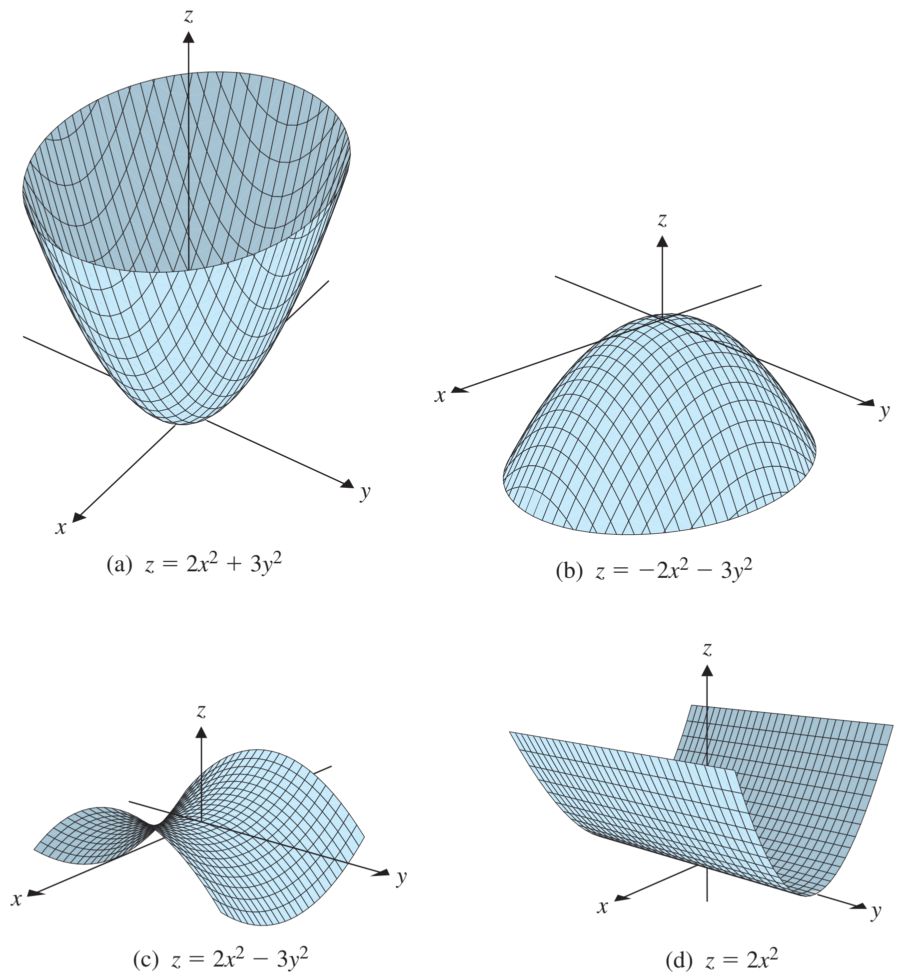


Figura 5.12
Gráficas de formas cuadráticas $f(x, y)$

Lo que hace a este tipo de análisis muy sencillo es el hecho de que estas formas cuadráticas no tienen términos en xy. La matriz asociada con tal forma cuadrática es una matriz diagonal. Por ejemplo,

$$2x^2 - 3y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

En general, la matriz de una forma cuadrática es una matriz simétrica y en la sección 5.4 se vio que tales matrices siempre pueden diagonalizarse. Ahora se usará este hecho para demostrar que, para *toda* forma cuadrática, es posible eliminar los términos en xy mediante un cambio adecuado de variable.

Sea $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ una forma cuadrática con n variables, con A como matriz simétrica de $n \times n$. Por el Teorema espectral, existe una matriz ortogonal Q que diagonaliza A; esto es, $Q^TAQ = D$, donde D es una matriz diagonal que muestra los eigenvalores de A. Ahora se establece

$$\mathbf{x} = Q\mathbf{y} \quad \text{o de manera equivalente,} \quad \mathbf{y} = Q^{-1}\mathbf{x} = Q^T\mathbf{x}$$

La sustitución en la forma cuadrática produce

$$\begin{aligned} \mathbf{x}^T A\mathbf{x} &= (Q\mathbf{y})^T A(Q\mathbf{y}) \\ &= \mathbf{y}^T Q^T A Q\mathbf{y} \\ &= \mathbf{y}^T D\mathbf{y} \end{aligned}$$

que es una forma cuadrática sin términos en xy, pues D es diagonal. Más aún, si los eigenvalores de A son $\lambda_1, \ldots, \lambda_n$, entonces Q puede elegirse de tal modo que

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Si $\mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^T$, entonces, con respecto a estas nuevas variables, la forma cuadrática se convierte en

$$\mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

Este proceso se llama ***diagonalización de una forma cuadrática***. Acaba de demostrarse el siguiente teorema, conocido como ***teorema de los ejes principales***. (La razón de este nombre se volverá clara en la siguiente subsección.)

Teorema 5.23 **El teorema de los ejes principales**

Toda forma cuadrática puede diagonalizarse. Específicamente, si A es la matriz simétrica de $n \times n$ asociada con la forma cuadrática $\mathbf{x}^T A\mathbf{x}$ y si Q es una matriz ortogonal tal que $Q^TAQ = D$ es una matriz diagonal, entonces el cambio de variable $\mathbf{x} = Q\mathbf{y}$ transforma la forma cuadrática $\mathbf{x}^T A\mathbf{x}$ en la forma cuadrática $\mathbf{y}^T D\mathbf{y}$, que no tiene términos en xy. Si los eigenvalores de A son $\lambda_1, \ldots, \lambda_n$ y $\mathbf{y} = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix}^T$, entonces

$$\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$$

Ejemplo 5.27 Encuentre un cambio de variable que transforme la forma cuadrática

$$f(x_1, x_2) = 5x_1^2 + 4x_1x_2 + 2x_2^2$$

en una sin términos en xy.

Solución La matriz de f es

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}$$

con eigenvalores $\lambda_1 = 6$ y $\lambda_2 = 1$. Los correspondientes eigenvectores unitarios son

$$\mathbf{q}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \quad \text{y} \quad \mathbf{q}_2 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$



(Compruebe esto.) Si se establece

$$Q = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}$$

entonces $Q^TAQ = D$. El cambio de variable $\mathbf{x} = Q\mathbf{y}$, donde

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \text{y} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

convierte f en

$$f(\mathbf{y}) = f(y_1, y_2) = [y_1 \quad y_2]\begin{bmatrix} 6 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 6y_1^2 + y_2^2$$



La forma cuadrática original $\mathbf{x}^TA\mathbf{x}$ y la ahora nueva $\mathbf{y}^TD\mathbf{y}$ (a la que se hace referencia en el teorema de los ejes principales) son *iguales* en el siguiente sentido. En el ejemplo 5.27, suponga que se quiere evaluar $f(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$ en $\mathbf{x} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$. Se tiene

$$f(-1, 3) = 5(-1)^2 + 4(-1)(3) + 2(3)^2 = 11$$

En términos de las nuevas variables,

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \mathbf{y} = Q^T\mathbf{x} = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ -7/\sqrt{5} \end{bmatrix}$$

de modo que

$$f(y_1, y_2) = 6y_1^2 + y_2^2 = 6(1/\sqrt{5})^2 + (-7/\sqrt{5})^2 = 55/5 = 11$$

exactamente como antes.

El teorema de los ejes principales tiene algunas consecuencias interesantes e importantes. Se considerarán dos de ellas. La primera se relaciona con los posibles *valores* que puede tomar una forma cuadrática.

Definición Una forma cuadrática $f(\mathbf{x}) = \mathbf{x}^TA\mathbf{x}$ se clasifica como una de las siguientes:

1. ***definida positiva*** si $f(\mathbf{x}) > 0$ para toda $\mathbf{x} \neq \mathbf{0}$
2. ***semidefinida positiva*** si $f(\mathbf{x}) \geq 0$ para toda $\mathbf{x}$
3. ***definida negativa*** si $f(\mathbf{x}) < 0$ para toda $\mathbf{x} \neq \mathbf{0}$
4. ***semidefinida negativa*** si $f(\mathbf{x}) \leq 0$ para toda $\mathbf{x}$
5. ***indefinida*** si $f(\mathbf{x})$ toma valores tanto positivos como negativos

Una matriz simétrica A se llama ***definida positiva***, ***semidefinida positiva***, ***definida negativa***, ***semidefinida negativa*** o ***indefinida*** si la forma cuadrática asociada $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ tiene la propiedad correspondiente.

Las formas cuadráticas de los incisos (a), (b), (c) y (d) de la figura 5.12 son definida positiva, definida negativa, indefinida y semidefinida positiva, respectivamente. El Teorema de los ejes principales facilita decir si una forma cuadrática tiene una de dichas propiedades.

Teorema 5.24 Sea A una matriz simétrica de $n \times n$. La forma cuadrática $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ es

a. positiva definida si y sólo si todos los eigenvalores de A son positivos.
b. semidefinida positiva si y sólo si todos los eigenvalores de A son no negativos.
c. definida negativa si y sólo si todos los eigenvalores de A son negativos.
d. semidefinida negativa si y sólo si todos los eigenvalores de A son no positivos.
e. indefinida si y sólo si A tiene eigenvalores tanto positivos como negativos.

En el ejercicio 49 se le pide demostrar el Teorema 5.24.

Ejemplo 5.28 Clasifique $f(x, y, z) = 3x^2 + 3y^2 + 3z^2 - 2xy - 2xz - 2yz$ como definida positiva, definida negativa, indefinida o ninguna de ellas.

Solución La matriz asociada con f es

$$\begin{bmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{bmatrix}$$

que tiene eigenvalores 1, 4, y 4. (Verifique esto.) Dado que todos estos eigenvalores son positivos, f es una forma cuadrática definida positiva.

Si una forma cuadrática $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ es definida positiva, entonces, dado que $f(\mathbf{0}) = 0$, el valor *mínimo* de $f(\mathbf{x})$ es 0 y se ubica en el origen. De igual modo, una forma cuadrática definida negativa tiene un máximo en el origen. Por ende, el Teorema 5.24 permite resolver con facilidad ciertos tipos de problemas de máximos/mínimos, sin depender del cálculo. Un tipo de problema que se incluye en esta categoría es el ***problema de optimización restringida***.

Con frecuencia es importante conocer los valores máximo o mínimo de una forma cuadrática sujeta a ciertas restricciones. (Tales problemas surgen no sólo en matemáticas, sino también en estadística, física, ingeniería y economía.) Se está interesado en encontrar los valores extremos de $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ sujetos a la restricción de que $\|\mathbf{x}\| = 1$. En el caso de una forma cuadrática con dos variables, puede visualizar qué significa el problema. La gráfica de $z = f(x, y)$ es una superficie en $\mathbb{R}^3$, y la restricción $\|\mathbf{x}\| = 1$ restringe el punto (x, y) a la circunferencia unitaria en el plano xy. Por tanto, se consideran aquellos puntos que están simultáneamente en la superficie y en el cilindro unitario perpendicular al plano xy. Dichos puntos forman una curva que se encuentra sobre la superficie, y se quieren los valores más alto y más bajo sobre esta curva. La figura 5.13 muestra esta situación para la forma cuadrática y la superficie correspondiente en la figura 5.12(c).

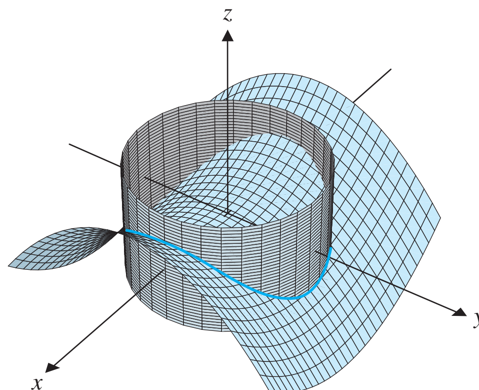


Figura 5.13
La intersección de $z = 2x^2 - 3y^2$ con el cilindro $x^2 + y^2 = 1$

En este caso, los valores máximo y mínimo de $f(x, y) = 2x^2 - 3y^2$ (los valores más alto y más bajo sobre la curva de intersección) son 2 y -3, respectivamente, que son justo los eigenvalores de la matriz asociada. El Teorema 5.25 muestra que este siempre es el caso.

Teorema 5.25

Sea $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ una forma cuadrática con matriz simétrica asociada A de $n \times n$. Sean $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ los eigenvalores de A. Entonces lo siguiente es verdadero, sujeto a la restricción $\|\mathbf{x}\| = 1$:

a. $\lambda_1 \geq f(\mathbf{x}) \geq \lambda_n$.
b. El valor máximo de $f(\mathbf{x})$ es λ_1 y ocurre cuando $\mathbf{x}$ es un eigenvector unitario correspondiente a λ_1.
c. El valor mínimo de $f(\mathbf{x})$ es λ_n y ocurre cuando $\mathbf{x}$ es un eigenvector unitario correspondiente a λ_n.

Demostración Como siempre, comience por diagonalizar ortogonalmente A. En concordancia, sea Q una matriz ortogonal tal que Q^TAQ es la matriz diagonal

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

Entonces, por el teorema de los ejes principales, el cambio de variable $\mathbf{x} = Q\mathbf{y}$ produce $\mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\mathbf{y}$. Ahora note que $\mathbf{y} = Q^T\mathbf{x}$ implica que

$$\mathbf{y}^T\mathbf{y} = (Q^T\mathbf{x})^T(Q^T\mathbf{x}) = \mathbf{x}^T(Q^T)^TQ^T\mathbf{x} = \mathbf{x}^TQQ^T\mathbf{x} = \mathbf{x}^T\mathbf{x}$$

pues $Q^T = Q^{-1}$. En consecuencia, al usar $\mathbf{x}\cdot\mathbf{x} = \mathbf{x}^T\mathbf{x}$, se ve que $\|\mathbf{y}\| = \sqrt{\mathbf{y}^T\mathbf{y}} = \sqrt{\mathbf{x}^T\mathbf{x}} = \|\mathbf{x}\| = 1$. Por tanto, si $\mathbf{x}$ es un vector unitario, también lo es su correspondiente $\mathbf{y}$, y los valores de $\mathbf{x}^TA\mathbf{x}$ y $\mathbf{y}^TD\mathbf{y}$ son los mismos.

(a) Para demostrar la propiedad (a), observe que si $\mathbf{y} = [y_1 \ \cdots \ y_n]^T$, entonces

$$\begin{aligned} f(\mathbf{x}) &= \mathbf{x}^T A\mathbf{x} = \mathbf{y}^T D\,\mathbf{y} \\ &= \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \\ &\le \lambda_1 y_1^2 + \lambda_1 y_2^2 + \cdots + \lambda_1 y_n^2 \\ &= \lambda_1(y_1^2 + y_2^2 + \cdots + y_n^2) \\ &= \lambda_1 \|\mathbf{y}\|^2 \\ &= \lambda_1 \end{aligned}$$

Por tanto, $f(\mathbf{x}) \le \lambda_1$ para todo $\mathbf{x}$ tal que $\|\mathbf{x}\| = 1$. La demostración de que $f(\mathbf{x}) \ge \lambda_n$ es similar. (Vea el ejercicio 59.)

(b) Si $\mathbf{q}_1$ es un eigenvector unitario correspondiente a λ_1, entonces $A\mathbf{q}_1 = \lambda_1\mathbf{q}_1$ y

$$f(\mathbf{q}_1) = \mathbf{q}_1^T A\mathbf{q}_1 = \mathbf{q}_1^T \lambda_1 \mathbf{q}_1 = \lambda_1(\mathbf{q}_1^T\mathbf{q}_1) = \lambda_1$$

Esto muestra que la forma cuadrática en realidad toma el valor λ_1 y, de este modo, por la propiedad (a), es el valor máximo de $f(\mathbf{x})$ y ocurre cuando $\mathbf{x} = \mathbf{q}_1$.

(c) En el ejercicio 60 se le pide demostrar esta propiedad.

Ejemplo 5.29

Encuentre los valores máximo y mínimo de la forma cuadrática $f(x_1, x_2) = 5x_1^2 + 4x_1x_2 + 2x_2^2$ sujeta a la restricción $x_1^2 + x_2^2 = 1$, y determine los valores x_1 y x_2 para los cuales se presentan cada uno de ellos.

Solución En el ejemplo 5.27 se encontró que f tiene los eigenvalores asociados $\lambda_1 = 6$ y $\lambda_2 = 1$, con correspondientes eigenvectores unitarios

$$\mathbf{q}_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} \quad \text{y} \quad \mathbf{q}_2 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

Por tanto, el valor máximo de f es 6 cuando $x_1 = 2/\sqrt{5}$ y $x_2 = 1/\sqrt{5}$. El valor mínimo de f es 1 cuando $x_1 = 1/\sqrt{5}$ y $x_2 = -2/\sqrt{5}$. (Observe que estos valores extremos ocurren dos veces, en direcciones opuestas, pues $-\mathbf{q}_1$ y $-\mathbf{q}_2$ también son eigenvectores unitarios para λ_1 y λ_2, respectivamente.)

Graficación de ecuaciones cuadráticas

La forma general de una ecuación cuadrática con dos variables x y y es

$$ax^2 + by^2 + cxy + dx + ey + f = 0$$

donde al menos uno de a, b y c es distinto de cero. Las gráficas de tales ecuaciones cuadráticas se llaman ***secciones cónicas*** (o ***cónicas***), pues pueden obtenerse al tomar secciones transversales de un cono (doble) (es decir, rebanándolo con un plano). Las más importantes de las secciones cónicas son las elipses (con las circunferencias como un caso especial), hipérbolas y parábolas. A las mismas se les llama cónicas ***no degeneradas***. La figura 5.14 muestra cómo surgen.

También es posible que una sección transversal de un cono resulte en un solo punto, una recta o un par de rectas. A ellas se les llaman cónicas ***degeneradas***. (Vea los ejercicios 81-86.)

Se dice que la gráfica de una cónica no degenerada está en ***posición estándar*** en relación con el eje coordenado si su ecuación puede expresarse en una de las formas de la figura 5.15.

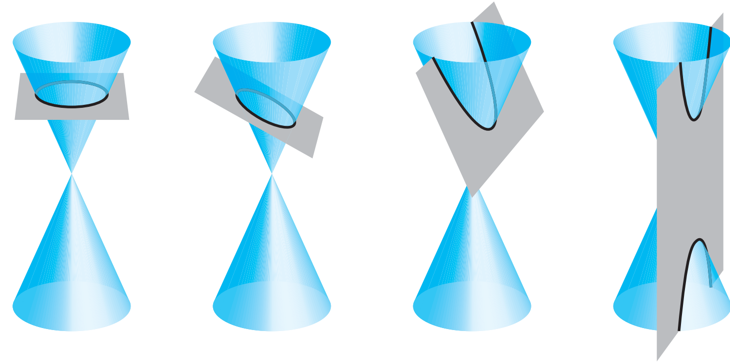
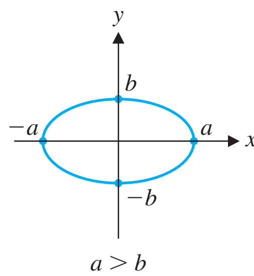
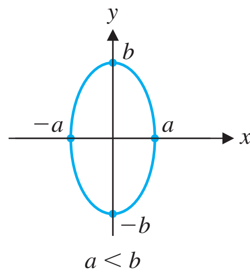


Figura 5.14
Las cónicas no degeneradas

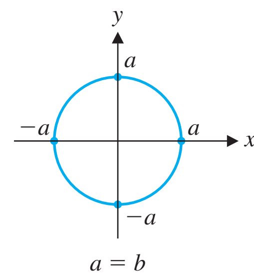
Elipse o circunferencia: $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1;\ a, b > 0$



$a > b$

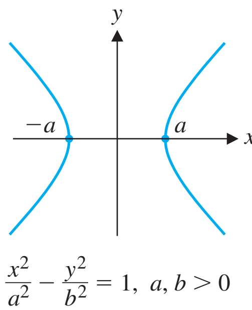


$a < b$

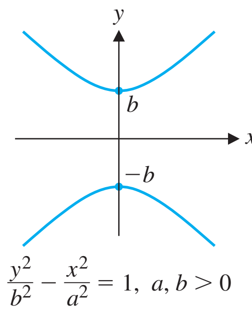


$a = b$

Hipérbola

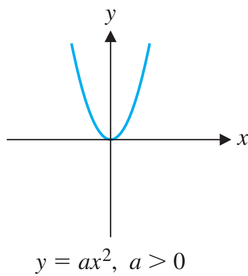


$\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1,\ a, b > 0$

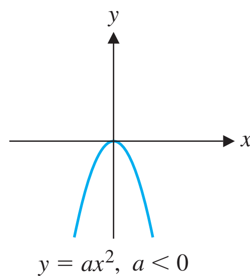


$\dfrac{y^2}{b^2} - \dfrac{x^2}{a^2} = 1,\ a, b > 0$

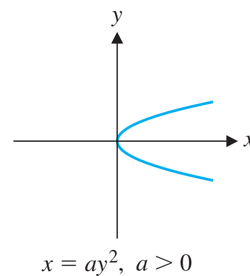
Parábola



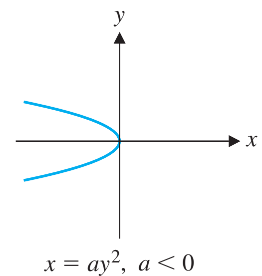
$y = ax^2,\ a > 0$



$y = ax^2,\ a < 0$



$x = ay^2,\ a > 0$



$x = ay^2,\ a < 0$

Figura 5.15
Cónicas no degeneradas en posición estándar

Ejemplo 5.30

Si es posible, escriba cada una de las siguientes ecuaciones cuadráticas en la forma de una cónica en posición estándar e identifique la gráfica resultante.

(a) $4x^2 + 9y^2 = 36$ (b) $4x^2 - 9y^2 + 1 = 0$ (c) $4x^2 - 9y = 0$

Solución (a) La ecuación $4x^2 + 9y^2 = 36$ puede escribirse en la forma

$$\frac{x^2}{9} + \frac{y^2}{4} = 1$$

de modo que su gráfica es una elipse que interseca el eje x en $(\pm 3, 0)$ y el eje y en $(0, \pm 2)$.

(b) La ecuación $4x^2 - 9y^2 + 1 = 0$ puede escribirse en la forma

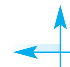
$$\frac{y^2}{\frac{1}{9}} - \frac{x^2}{\frac{1}{4}} = 1$$

de modo que su gráfica es una hipérbola que se abre hacia arriba y abajo, e interseca el eje y en $(0, \pm\frac{1}{3})$.

(c) La ecuación $4x^2 - 9y = 0$ puede escribirse en la forma

$$y = \frac{4}{9}x^2$$

de modo que su gráfica es una parábola que abre hacia arriba.



Si una ecuación cuadrática contiene demasiados términos para escribirse en una de las formas de la figura 5.15, entonces su gráfica no está en posición estándar. Cuando existen términos adicionales más no término xy, la gráfica de la cónica se *trasladó* fuera de la posición estándar.

Ejemplo 5.31

Identifique y grafique la cónica cuya ecuación es

$$x^2 + 2y^2 - 6x + 8y + 9 = 0$$

Solución Comience por agrupar los términos x y y por separado para obtener

$$(x^2 - 6x) + (2y^2 + 8y) = -9$$

o

$$(x^2 - 6x) + 2(y^2 + 4y) = -9$$

A continuación, complete los cuadrados en las dos expresiones entre paréntesis para obtener

$$(x^2 - 6x + 9) + 2(y^2 + 4y + 4) = -9 + 9 + 8$$

o

$$(x - 3)^2 + 2(y + 2)^2 = 8$$

Ahora haga las sustituciones $x' = x - 3$ y $y' = y + 2$, y convierta la ecuación anterior en

$$(x')^2 + 2(y')^2 = 8 \quad \text{o} \quad \frac{(x')^2}{8} + \frac{(y')^2}{4} = 1$$

Esta es la ecuación de una elipse en posición estándar en el sistema coordenado $x'y'$, que interseca el eje x' en $(\pm 2\sqrt{2}, 0)$ y el eje y en $(0, \pm 2)$. El origen en el sistema coordenado $x'y'$ está en $x = 3, y = -2$, de modo que la elipse se trasladó fuera de la posición estándar 3 unidades hacia la derecha y 2 unidades hacia abajo. Su gráfica se muestra en la figura 5.16.

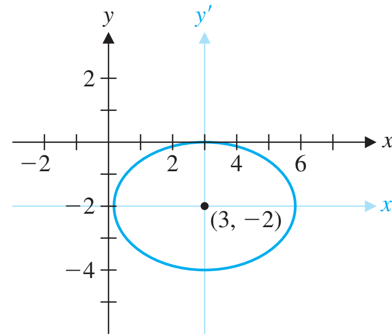


Figura 5.16
Una elipse trasladada

Si una ecuación cuadrática contiene un término en xy, entonces representa una cónica que se *rotó*.

Ejemplo 5.32

Identifique y grafique de la cónica cuya ecuación es

$$5x^2 + 4xy + 2y^2 = 6$$

Solución El lado izquierdo de la ecuación es una forma cuadrática, de modo que puede escribirla en forma matricial como $\mathbf{x}^T A\mathbf{x} = 6$, donde

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix}$$

En el ejemplo 5.27 se encontró que los eigenvalores de A son 6 y 1, y una matriz Q que diagonaliza ortogonalmente a A es

$$Q = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}$$

Observe que det $Q = -1$. En este ejemplo se intercambiarán las columnas de esta matriz para hacer el determinante igual a $+1$. Entonces Q será la matriz de una *rotación*, por el ejercicio 28 de la sección 5.1. Siempre es posible reordenar las columnas de una matriz ortogonal Q para hacer su determinante igual a $+1$. (¿Por qué?) En vez de ello, establezca

$$Q = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

de modo que

$$Q^T A Q = \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} = D$$

El cambio de variable $\mathbf{x} = Q\mathbf{x}'$ convierte la ecuación dada en la forma $(\mathbf{x}')^T D\mathbf{x}' = 6$ mediante una rotación. Si $\mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}$, entonces esta ecuación es justo

$$(x')^2 + 6(y')^2 = 6 \quad \text{o} \quad \frac{(x')^2}{6} + (y')^2 = 1$$

que representa una elipse en el sistema coordenado $x'y'$.

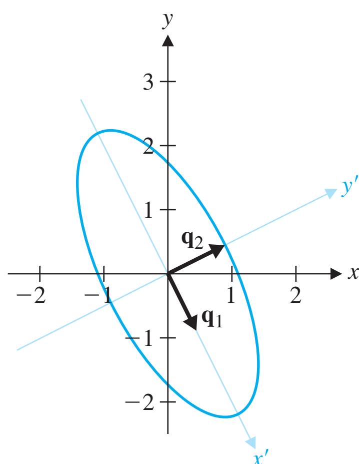


Figura 5.17
Una elipse rotada

Para graficar esta elipse, es necesario conocer cuáles vectores desempeñan el papel de $\mathbf{e}_1' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ y $\mathbf{e}_2' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ en el nuevo sistema coordenado. (Estos dos vectores localizan las posiciones de los ejes x' y y'.) Pero, a partir de $\mathbf{x} = Q\mathbf{x}'$, se tiene

$$Q\mathbf{e}_1' = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

y

$$Q\mathbf{e}_2' = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

Éstas son las columnas $\mathbf{q}_1$ y $\mathbf{q}_2$ de Q, ¡que son los eigenvectores de A! El hecho de que estos vectores sean ortonormales concuerda perfectamente con el hecho de que el cambio de variable es justo una rotación. La gráfica se muestra en la figura 5.17.

Ahora puede ver por qué se llama así al teorema de los ejes principales. Si una matriz simétrica real A surge como la matriz de coeficientes de una ecuación cuadrática, los eigenvectores de A ofrecen las direcciones de los ejes principales de la gráfica correspondiente.

Es posible que la gráfica de una cónica esté tanto rotada como trasladada fuera de la posición estándar, como se ilustra en el ejemplo 5.33.

Ejemplo 5.33

Identifique y grafique la cónica cuya ecuación es

$$5x^2 + 4xy + 2y^2 - \frac{28}{\sqrt{5}}x - \frac{4}{\sqrt{5}}y + 4 = 0$$

Solución La estrategia es eliminar primero el término en xy. En forma matricial, la ecuación es $\mathbf{x}^T A\mathbf{x} + B\mathbf{x} + 4 = 0$, donde

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 2 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -\frac{28}{\sqrt{5}} & -\frac{4}{\sqrt{5}} \end{bmatrix}$$

El término en xy proviene de la forma cuadrática $\mathbf{x}^T A\mathbf{x}$, que se diagonaliza como en el ejemplo 5.32 al establecer $\mathbf{x} = Q\mathbf{x}'$, donde

$$Q = \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

Entonces, como en el ejemplo 5.32,

$$\mathbf{x}^T A\mathbf{x} = (\mathbf{x}')^T D\mathbf{x}' = (x')^2 + 6(y')^2$$

Pero ahora también se tiene

$$B\mathbf{x} = BQ\mathbf{x}' = \begin{bmatrix} -\frac{28}{\sqrt{5}} & -\frac{4}{\sqrt{5}} \end{bmatrix}\begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}\begin{bmatrix} x' \\ y' \end{bmatrix} = -4x' - 12y'$$

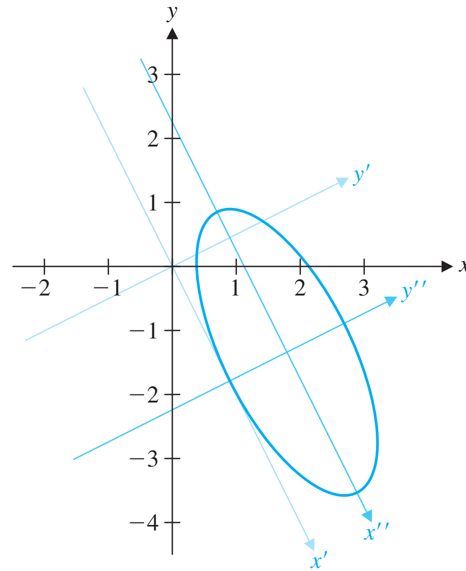


Figura 5.18

Por tanto, en términos de x' y y', la ecuación dada se convierte en

$$(x')^2 + 6(y')^2 - 4x' - 12y' + 4 = 0$$

Para llevar a la posición estándar la cónica representada por esta ecuación, es necesario *trasladar* los ejes $x'y'$. Esto se hace al completar los cuadrados, como en el ejemplo 5.31. Se tiene

$$((x')^2 - 4x' + 4) + 6((y')^2 - 2y' + 1) = -4 + 4 + 6 = 6$$

o

$$(x' - 2)^2 + 6(y' - 1)^2 = 6$$

Esto produce las ecuaciones de traslación

$$x'' = x' - 2 \quad \text{y} \quad y'' = y' - 1$$

En el sistema coordenado $x''y''$, la ecuación es simplemente

$$(x'')^2 + 6(y'')^2 = 6$$

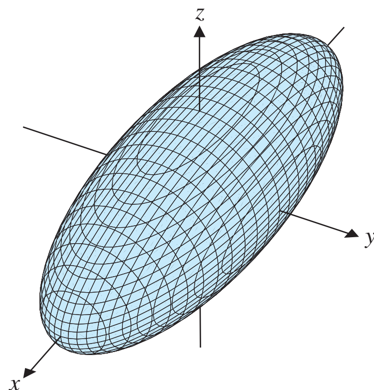
que es la ecuación de una elipse (como en el ejemplo 5.32). Esta elipse se puede bosquejar al rotar primero y luego trasladar. La gráfica resultante se muestra en la figura 5.18.

La forma general de una ecuación cuadrática con tres variables x, y y z es

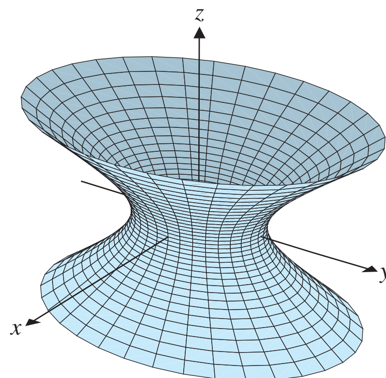
$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + iz + j = 0$$

donde al menos uno de $a, b, \ldots, f$ es distinto de cero. La gráfica de tal ecuación cuadrática se llama ***superficie cuádrica*** (o ***cuádrica***). Nuevamente, para reconocer una cuá-

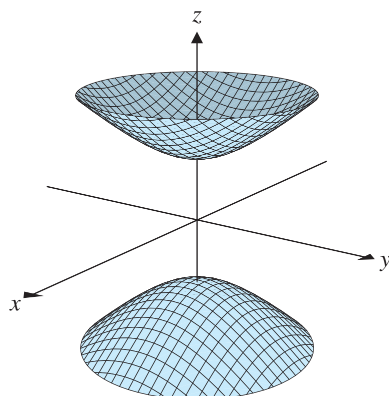
Elipsoide: $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$



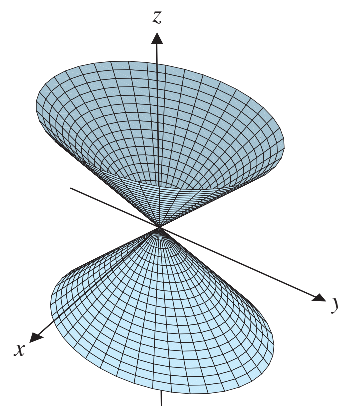
Hiperboloide de una hoja: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$



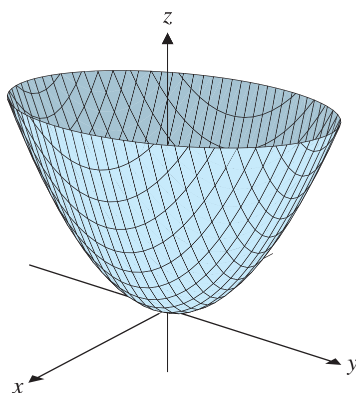
Hiperboloide de dos hojas: $\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = -1$



Cono elíptico: $z^2 = \frac{x^2}{a^2} + \frac{y^2}{b^2}$



Paraboloide elíptico: $z = \frac{x^2}{a^2} + \frac{y^2}{b^2}$



Paraboloide hiperbólico: $z = \frac{x^2}{a^2} - \frac{y^2}{b^2}$

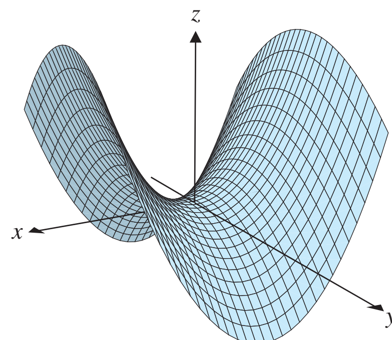


Figura 5.19
Superficies cuádricas

drica es necesario ponerla en posición estándar. En la figura 5.19 se muestran algunas cuádricas en posición estándar; otras se obtienen al permutar las variables.

Ejemplo 5.34 Identifique la superficie cuádrica cuya ecuación es

$$5x^2 + 11y^2 + 2z^2 + 16xy + 20xz - 4yz = 36$$

Solución La ecuación puede escribirse en forma matricial como $\mathbf{x}^T A\mathbf{x} = 36$, donde

$$A = \begin{bmatrix} 5 & 8 & 10 \\ 8 & 11 & -2 \\ 10 & -2 & 2 \end{bmatrix}$$

Se encuentra que los eigenvalores de A son 18, 9 y -9, con sus correspondientes eigenvectores ortogonales

$$\begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 \\ -1 \\ -2 \end{bmatrix}$$

respectivamente. Al normalizarlos se obtiene

$$\mathbf{q}_1 = \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \quad \text{y} \quad \mathbf{q}_3 = \begin{bmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

y forman la matriz ortogonal

$$Q = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \mathbf{q}_3] = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} & -\frac{2}{3} \end{bmatrix}$$

Note que, con la finalidad de que Q sea la matriz de una rotación, se requiere que det $Q = 1$, lo que es verdadero en este caso. (De otro modo, det $Q = -1$ y el intercambio de dos columnas modifica el signo del determinante.) Por tanto,

$$Q^T AQ = D = \begin{bmatrix} 18 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & -9 \end{bmatrix}$$

y, con el cambio de variable $\mathbf{x} = Q\mathbf{x}'$, se obtiene $\mathbf{x}^T A\mathbf{x} = (\mathbf{x}')D\mathbf{x}' = 36$, de modo que

$$18(x')^2 + 9(y')^2 - 9(z')^2 = 36 \quad \text{o} \quad \frac{(x')^2}{2} + \frac{(y')^2}{4} - \frac{(z')^2}{4} = 1$$

A partir de la figura 5.19, esta ecuación se reconoce como la de un hiperboloide de una hoja. Los ejes x', y' y z' están en las direcciones de los eigenvectores $\mathbf{q}_1$, $\mathbf{q}_2$ y $\mathbf{q}_3$, respectivamente. La gráfica se muestra en la figura 5.20.

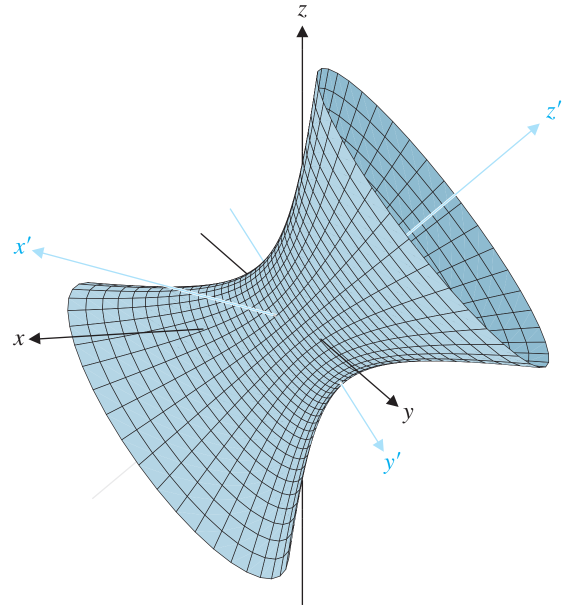


Figura 5.20
Un hiperboloide de una hoja en posición no estándar

También puede identificar y graficar cuádricas que se hayan trasladado fuera de la posición estándar usando el "método de completar cuadrados" de los ejemplos 5.31 y 5.33. Se le pedirá hacerlo en los ejercicios.

Ejercicios 5.5

Códigos duales

En los ejercicios 1-4, G es una matriz generadora para un código C. Lleve G a forma estándar y determine si el código correspondiente es igual a C.

1. $G = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}$

2. $G = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

3. $G = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

4. $G = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}$

En los ejercicios 5-8, P es una matriz de control de paridad para un código C. Lleve P a forma estándar y determine si el código correspondiente es igual a C.

5. $P = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$

6. $P = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

7. $P = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$

8. $P = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$

En los ejercicios 9-12, encuentre el código dual $C^{\perp}$ del código C.

9. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

10. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

11. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

12. $C = \left\{ \begin{bmatrix} 0\\0\\0\\0\\0 \end{bmatrix}, \begin{bmatrix} 0\\1\\1\\0\\1 \end{bmatrix}, \begin{bmatrix} 1\\0\\0\\1\\0 \end{bmatrix}, \begin{bmatrix} 1\\1\\1\\1\\1 \end{bmatrix} \right\}$

En los ejercicios 13-16, se proporciona una matriz generadora G o una matriz de verificación de paridad P para un código C. Encuentre una matriz generadora $G^{\perp}$ y una matriz de control de paridad $P^{\perp}$ para el código dual de C.

13. $G = \begin{bmatrix} 1 & 1\\ 1 & 1\\ 1 & 0\\ 0 & 1 \end{bmatrix}$ 14. $G = \begin{bmatrix} 1 & 0\\ 0 & 1\\ 1 & 0\\ 1 & 1\\ 0 & 1 \end{bmatrix}$

15. $P = \begin{bmatrix} 1 & 1 & 1 & 0\\ 0 & 1 & 0 & 1 \end{bmatrix}$ 16. $P = \begin{bmatrix} 1 & 0 & 1 & 1 & 0\\ 0 & 1 & 0 & 0 & 1\\ 0 & 0 & 1 & 0 & 1 \end{bmatrix}$

17. Encuentre las matrices generadora y de control de paridad para el dual del código de Hamming (7, 4) del ejemplo 3.71.

El ***código de paridad par*** *E_n es el subconjunto de $\mathbb{Z}_2^n$ que consiste de todos los vectores con peso par. El* ***código de repetición n-veces*** *Rep_n es el subconjunto de $\mathbb{Z}_2^n$ que consiste sólo de dos vectores* **0** *y* **1** *(todos ceros y todos 1, respectivamente).*

18. **(a)** Encuentre las matrices generadora y de verificación de paridad para E_3 y Rep_3.
(b) Demuestre que E_3 y Rep_3 son duales mutuamente.

19. Demuestre que E_n y Rep_n son duales mutuamente.

20. Si C y D son códigos y $C \subseteq D$, demuestre que $D^{\perp} \subseteq C^{\perp}$.

21. Demuestre que, si C es un código con una matriz generadora, entonces $(C^{\perp})^{\perp} = C$.

22. Encuentre un código autodual de longitud 6.

Formas cuadráticas

En los ejercicios 23-28, evalúe la forma cuadrática $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ para los A y **x** *dados.*

23. $A = \begin{bmatrix} 2 & 3\\ 3 & 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x\\ y \end{bmatrix}$

24. $A = \begin{bmatrix} 5 & 1\\ 1 & -1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1\\ x_2 \end{bmatrix}$

25. $A = \begin{bmatrix} 3 & -2\\ -2 & 4 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1\\ 6 \end{bmatrix}$

26. $A = \begin{bmatrix} 1 & 0 & -3\\ 0 & 2 & 1\\ -3 & 1 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x\\ y\\ z \end{bmatrix}$

27. $A = \begin{bmatrix} 1 & 0 & -3\\ 0 & 2 & 1\\ -3 & 1 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 2\\ -1\\ 1 \end{bmatrix}$

28. $A = \begin{bmatrix} 2 & 2 & 0\\ 2 & 0 & 1\\ 0 & 1 & 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} 1\\ 2\\ 3 \end{bmatrix}$

En los ejercicios 29-34, encuentre la matriz simétrica A asociada con la forma cuadrática dada.

29. $x_1^2 + 2x_2^2 + 6x_1x_2$ 30. x_1x_2

31. $3x^2 - 3xy - y^2$ 32. $x_1^2 - x_3^2 + 8x_1x_2 - 6x_2x_3$

33. $5x_1^2 - x_2^2 + 2x_3^2 + 2x_1x_2 - 4x_1x_3 + 4x_2x_3$

34. $2x^2 - 3y^2 + z^2 - 4xz$

Diagonalice las formas cuadráticas de los ejercicios 35-40 al encontrar una matriz ortogonal Q tal que el cambio de variable **x** $= Q$**y** *transforme la forma dada en una sin términos xy. Proporcione Q y la nueva forma cuadrática.*

35. $2x_1^2 + 5x_2^2 - 4x_1x_2$ 36. $x^2 + 8xy + y^2$

37. $7x_1^2 + x_2^2 + x_3^2 + 8x_1x_2 + 8x_1x_3 - 16x_2x_3$

38. $x_1^2 + x_2^2 + 3x_3^2 - 4x_1x_2$

39. $x^2 + z^2 - 2xy + 2yz$

40. $2xy + 2xz + 2yz$

Clasifique cada una de las formas cuadráticas de los ejercicios 41-48 como definida positiva, semidefinida positiva, definida negativa, semidefinida negativa o indefinida.

41. $x_1^2 + 2x_2^2$ 42. $x_1^2 + x_2^2 - 2x_1x_2$

43. $-2x^2 - 2y^2 + 2xy$ 44. $x^2 + y^2 + 4xy$

45. $2x_1^2 + 2x_2^2 + 2x_3^2 + 2x_1x_2 + 2x_1x_3 + 2x_2x_3$

46. $x_1^2 + x_2^2 + x_3^2 + 2x_1x_3$ 47. $x_1^2 + x_2^2 - x_3^2 + 4x_1x_2$

48. $-x^2 - y^2 - z^2 - 2xy - 2xz - 2yz$

49. Demuestre el Teorema 5.24.

50. Sea $A = \begin{bmatrix} a & b\\ b & d \end{bmatrix}$ una matriz simétrica de 2×2.

Demuestre que A es definida positiva si y sólo si $a > 0$ y $\det A > 0$. [*Sugerencia:* $ax^2 + 2bxy + dy^2 =$

$$a\left(x + \frac{b}{a}y\right)^2 + \left(d - \frac{b^2}{a}\right)y^2.]$$

51. Sea B una matriz invertible. Demuestre que $A = B^TB$ es definida positiva.

52. Sea A una matriz simétrica definida positiva. Demuestre que existe una matriz invertible B tal que $A = B^TB$. [*Sugerencia:* use el teorema espectral para escribir $A = QDQ^T$. Luego demuestre que D puede factorizarse como C^TC para alguna matriz invertible C.]

53. Sean A y B matrices simétricas definidas positivas de $n \times n$, y sea c un escalar positivo. Demuestre que las siguientes matrices son definidas positivas.

(a) cA **(b)** A^2 **(c)** $A + B$

(d) A^{-1} (Primero demuestre que A necesariamente es invertible.)

54. Sea A una matriz simétrica definida positiva. Demuestre que existe una matriz simétrica definida positiva B tal que $A = B^2$. (Tal matriz B se llama ***raíz cuadrada*** de A.)

En los ejercicios 55-58, encuentre los valores máximo y mínimo de la forma cuadrática $f(\mathbf{x})$ en el ejercicio dado, sujeta a la restricción $\|\mathbf{x}\| = 1$ y determine los valores de $\mathbf{x}$ para los cuales esto ocurre.

55. Ejercicio 42 **56.** Ejercicio 44

57. Ejercicio 45 **58.** Ejercicio 46

59. Termine la demostración del Teorema 5.25(a).

60. Demuestre el Teorema 5.25(c).

Graficación de ecuaciones cuadráticas

En los ejercicios 61-66, identifique la gráfica de la ecuación dada.

61. $x^2 + 5y^2 = 25$ **62.** $x^2 - y^2 - 4 = 0$

63. $x^2 - y - 1 = 0$ **64.** $2x^2 + y^2 - 8 = 0$

65. $3x^2 = y^2 - 1$ **66.** $x = -2y^2$

En los ejercicios 67-72, use una traslación de ejes para poner la cónica en posición estándar. Identifique la gráfica, proporcione su ecuación en el sistema coordenado trasladado y bosqueje la curva.

67. $x^2 + y^2 - 4x - 4y + 4 = 0$

68. $4x^2 + 2y^2 - 8x + 12y + 6 = 0$

69. $9x^2 - 4y^2 - 4y = 37$ **70.** $x^2 + 10x - 3y = -13$

71. $2y^2 + 4x + 8y = 0$

72. $2y^2 - 3x^2 - 18x - 20y + 11 = 0$

En los ejercicios 73-76, use una rotación de ejes para poner la cónica en posición estándar. Identifique la gráfica, proporcione su ecuación en el sistema coordenado rotado y bosqueje la curva.

73. $x^2 + xy + y^2 = 6$ **74.** $4x^2 + 10xy + 4y^2 = 9$

75. $4x^2 + 6xy - 4y^2 = 5$ **76.** $3x^2 - 2xy + 3y^2 = 8$

En los ejercicios 77-80, identifique la cónica con la ecuación dada y proporcione su ecuación en forma estándar.

77. $3x^2 - 4xy + 3y^2 - 28\sqrt{2}x + 22\sqrt{2}y + 84 = 0$

78. $6x^2 - 4xy + 9y^2 - 20x - 10y - 5 = 0$

79. $2xy + 2\sqrt{2}x - 1 = 0$

80. $x^2 - 2xy + y^2 + 4\sqrt{2}x - 4 = 0$

*En ocasiones, la gráfica de una ecuación cuadrática es una recta, un par de rectas o un solo punto. A tales gráficas se les conoce como **cónica degenerada**. También es posible que la ecuación no se satisfaga con valor alguno de las variables, en cuyo caso no existe gráfica en absoluto y a la cónica se le conoce como **cónica imaginaria**. En los ejercicios 81-86, identifique la cónica con la ecuación dada o como degenerada o como imaginaria y, donde sea posible, bosqueje la gráfica.*

81. $x^2 - y^2 = 0$ **82.** $x^2 + 2y^2 + 2 = 0$

83. $3x^2 + y^2 = 0$ **84.** $x^2 + 2xy + y^2 = 0$

85. $x^2 - 2xy + y^2 + 2\sqrt{2}x - 2\sqrt{2}y = 0$

86. $2x^2 + 2xy + 2y^2 + 2\sqrt{2}x - 2\sqrt{2}y + 6 = 0$

87. Sea A una matriz simétrica de 2×2 y sea k un escalar. Demuestre que la gráfica de la ecuación cuadrática $\mathbf{x}^TA\mathbf{x} = k$ es

(a) una hipérbola si $k \neq 0$ y $\det A < 0$

(b) una elipse, circunferencia o cónica imaginaria si $k \neq 0$ y $\det A > 0$

(c) un par de rectas o una cónica imaginaria si $k \neq 0$ y $\det A = 0$

(d) un par de rectas o un solo punto si $k = 0$ y $\det A \neq 0$

(e) una recta si $k = 0$ y $\det A = 0$ [*Sugerencia:* use el teorema de ejes principales.]

En los ejercicios 88-95, identifique la cuádrica con la ecuación dada y proporcione su ecuación en forma estándar.

88. $4x^2 + 4y^2 + 4z^2 + 4xy + 4xz + 4yz = 8$

89. $x^2 + y^2 + z^2 - 4yz = 1$

90. $-x^2 - y^2 - z^2 + 4xy + 4xz + 4yz = 12$

91. $2xy + z = 0$

92. $16x^2 + 100y^2 + 9z^2 - 24xz - 60x - 80z = 0$

93. $x^2 + y^2 - 2z^2 + 4xy - 2xz + 2yz - x + y + z = 0$

94. $10x^2 + 25y^2 + 10z^2 - 40xz + 20\sqrt{2}x + 50y + 20\sqrt{2}z = 15$

95. $11x^2 + 11y^2 + 14z^2 + 2xy + 8xz - 8yz - 12x + 12y + 12z = 6$

96. Sea A una matriz real de 2×2 con eigenvalores complejos $\lambda = a \pm bi$ tales que $b \neq 0$ y $|\lambda| = 1$. Demuestre que toda trayectoria del sistema dinámico $\mathbf{x}_{k+1} = A\mathbf{x}_k$ se encuentra sobre una elipse. [*Sugerencia:* el Teorema 4.43 muestra que si $\mathbf{v}$ es un eigenvector correspondiente a $\lambda = a - bi$, entonces la matriz $P = [\text{Re}\,\mathbf{v} \quad \text{Im}\,\mathbf{v}]$ es invertible y $A = P\begin{bmatrix} a & -b \\ b & a \end{bmatrix}P^{-1}$. Establezca $B = (PP^T)^{-1}$. Demuestre que la cuadrática $\mathbf{x}^T B\mathbf{x} = k$ define una elipse para toda $k > 0$ y pruebe que si $\mathbf{x}$ está sobre dicha elipse, también lo hace $A\mathbf{x}$.]

Repaso del capítulo

Definiciones y conceptos clave

base ortogonal, 381
base ortonormal, 383
complemento ortogonal de un subespacio, 389
conjunto ortogonal de vectores, 380
conjunto ortonormal de vectores, 383
descomposición espectral, 416
factorización QR, 404
matriz ortogonal, 385
matriz diagonalizable ortogonalmente, 411
proceso de Gram-Schmidt, 400
propiedades de las matrices ortogonales, 386-387
proyección ortogonal, 393
subespacios fundamentales de una matriz, 391
teorema de descomposición ortogonal, 395
teorema del rank, 397
teorema espectral, 414

Preguntas de repaso

1. Marque cada uno de los siguientes enunciados como verdadero o falso:

(a) Todo conjunto ortonormal de vectores es linealmente independiente.

(b) Todo subespacio distinto de cero de $\mathbb{R}^n$ tiene una base ortogonal.

(c) Si A es una matriz cuadrada con renglones ortonormales, entonces A es una matriz ortogonal.

(d) Toda matriz ortogonal es invertible.

(e) Si A es una matriz con $\det A = 1$, entonces A es una matriz ortogonal.

(f) Si A es una matriz de $m \times n$ tal que $(\text{renglón}(A))^{\perp} = \mathbb{R}^n$, entonces A debe ser la matriz cero.

(g) Si W es un subespacio de $\mathbb{R}^n$ y $\mathbf{v}$ es un vector en $\mathbb{R}^n$ tal que $\text{proy}_W(\mathbf{v}) = \mathbf{0}$, entonces $\mathbf{v}$ debe ser el vector cero.

(h) Si A es una matriz ortogonal simétrica, entonces $A^2 = I$.

(i) Toda matriz diagonalizable ortogonalmente es invertible.

(j) Dados cualesquiera números reales $n, \lambda_1, \ldots, \lambda_n$, existe una matriz simétrica de $n \times n$ con $\lambda_1, \ldots, \lambda_n$ como sus eigenvalores.

2. Encuentre todos los valores de a y b tales que

$$\left\{\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} a \\ b \\ 3 \end{bmatrix}\right\}$$ es un conjunto ortogonal de vectores.

3. Encuentre el vector coordenado $[\mathbf{v}]_{\mathcal{B}}$ de $\mathbf{v} = \begin{bmatrix} 7 \\ -3 \\ 2 \end{bmatrix}$ con respecto a la base ortogonal

$$\mathcal{B} = \left\{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}\right\} \text{ de } \mathbb{R}^3$$

4. El vector coordenado de un vector $\mathbf{v}$ con respecto a una base ortonormal $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2\}$ de $\mathbb{R}^2$ es $[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} -3 \\ 1/2 \end{bmatrix}$. Si $\mathbf{v}_1 = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$, encuentre todos los posibles vectores $\mathbf{v}$.

5. Demuestre que $\begin{bmatrix} 6/7 & 2/7 & 3/7 \\ -1/\sqrt{5} & 0 & 2/\sqrt{5} \\ 4/7\sqrt{5} & -15/7\sqrt{5} & 2/7\sqrt{5} \end{bmatrix}$ es una matriz ortogonal.

6. Si $\begin{bmatrix} 1/2 & a \\ b & c \end{bmatrix}$ es una matriz ortogonal, encuentre todos los posibles valores de a, b y c.

7. Si Q es una matriz ortogonal de $n \times n$ y $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es un conjunto ortonormal en $\mathbb{R}^n$, demuestre que $\{Q\mathbf{v}_1, \ldots, Q\mathbf{v}_k\}$ es un conjunto ortonormal.

8. Si Q es una matriz de $n \times n$ tal que los ángulos $\angle(Q\mathbf{x}, Q\mathbf{y})$ y $\angle(\mathbf{x}, \mathbf{y})$ son iguales para todos los vectores $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{R}^n$, pruebe que Q es una matriz ortogonal.

En las preguntas 9-12, encuentre una base para $W^\perp$.

9. W es la recta en $\mathbb{R}^2$ con ecuación general $2x - 5y = 0$

10. W es la recta en $\mathbb{R}^3$ con ecuaciones paramétricas

$$x = t$$
$$y = 2t$$
$$z = -t$$

11. $W = \text{gen}\left\{\begin{bmatrix}1\\-1\\4\end{bmatrix}, \begin{bmatrix}0\\1\\-3\end{bmatrix}\right\}$

12. $W = \text{gen}\left\{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\-1\\1\\2\end{bmatrix}\right\}$

13. Encuentre bases para cada uno de los cuatro subespacios fundamentales de

$$A = \begin{bmatrix}1 & -1 & 2 & 1 & 3\\ -1 & 2 & -2 & 1 & -2\\ 2 & 1 & 4 & 8 & 9\\ 3 & -5 & 6 & -1 & 7\end{bmatrix}$$

14. Encuentre la descomposición ortogonal de

$$\mathbf{v} = \begin{bmatrix}1\\0\\-1\\2\end{bmatrix}$$

con respecto a

$$W = \text{gen}\left\{\begin{bmatrix}0\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\1\\-1\end{bmatrix}, \begin{bmatrix}3\\1\\-2\\1\end{bmatrix}\right\}$$

15. **(a)** Aplique el Proceso de Gram-Schmidt a

$$\mathbf{x}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix}1\\1\\1\\0\end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix}$$

para encontrar una base ortogonal para $W = \text{gen}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$.

(b) Use el resultado del inciso (a) para encontrar una factorización QR de $A = \begin{bmatrix}1 & 1 & 0\\ 1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 0 & 1\end{bmatrix}$.

16. Encuentre una base ortogonal para $\mathbb{R}^4$ que contenga los vectores $\begin{bmatrix}1\\0\\2\\2\end{bmatrix}$ y $\begin{bmatrix}0\\1\\1\\-1\end{bmatrix}$

17. Encuentre una base ortogonal para el subespacio

$$W = \left\{\begin{bmatrix}x_1\\x_2\\x_3\\x_4\end{bmatrix} : x_1 + x_2 + x_3 + x_4 = 0\right\} \text{ de } \mathbb{R}^4.$$

18. Sea $A = \begin{bmatrix}2 & 1 & -1\\ 1 & 2 & 1\\ -1 & 1 & 2\end{bmatrix}$.

(a) Diagonalice ortogonalmente A.
(b) Proporcione la descomposición espectral de A.

19. Encuentre una matriz simétrica con eigenvalores $\lambda_1 = \lambda_2 = 1, \lambda_3 = -2$ y eigenespacios

$$E_1 = \text{gen}\left(\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right), E_{-2} = \text{gen}\left(\begin{bmatrix}1\\-1\\0\end{bmatrix}\right)$$

20. Si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base ortonormal para $\mathbb{R}^n$ y

$$A = c_1\mathbf{v}_1\mathbf{v}_1^T + c_2\mathbf{v}_2\mathbf{v}_2^T + \cdots + c_n\mathbf{v}_n\mathbf{v}_n^T$$

demuestre que A es una matriz simétrica con eigenvalores $c_1, c_2, \ldots, c_n$ y correspondientes eigenvectores $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$.

6

Espacios vectoriales

El álgebra es generosa; con frecuencia da más de lo que pide.
–Jean le Rond d'Alembert (1717-1783)

En Carl B. Boyer,
A History of Mathematics,
Wiley, 1968, p. 481

6.0 Introducción: Fibonacci en el espacio (vectorial)

La sucesión de Fibonacci se introdujo en la sección 4.6. Es la sucesión

$$0, 1, 1, 2, 3, 5, 8, 13, \ldots$$

de enteros no negativos con la propiedad de que, después de los primeros dos términos, cada término es la suma de los dos términos que lo anteceden. Por tanto, $0 + 1 = 1, 1 + 1 = 2, 1 + 2 = 3, 2 + 3 = 5$, etcétera.

Si los términos de la sucesión de Fibonacci se denotan por $f_0, f_1, f_2, \ldots$, entonces toda la sucesión está completamente determinada al especificar que

$$f_0 = 0, f_1 = 1 \quad \text{y} \quad f_n = f_{n-1} + f_{n-2} \quad \text{para } n \geq 2$$

Por analogía con la notación vectorial, se escribe una sucesión $x_0, x_1, x_2, x_3, \ldots$ como

$$\mathbf{x} = [x_0, x_1, x_2, x_3, \ldots)$$

Entonces la sucesión de Fibonacci se convierte en

$$\mathbf{f} = [f_0, f_1, f_2, f_3, \ldots) = [0, 1, 1, 2, \ldots)$$

Ahora se generaliza esta noción.

Definición Una ***sucesión tipo Fibonacci*** es *cualquier* sucesión $\mathbf{x} = [x_0, x_1, x_2, x_3, \ldots)$ tal que x_0 y x_1 son números reales y $x_n = x_{n-1} + x_{n-2}$ para $n \geq 2$.

Por ejemplo, $[1, \sqrt{2}, 1 + \sqrt{2}, 1 + 2\sqrt{2}, 2 + 3\sqrt{2}, \ldots)$ es una sucesión tipo Fibonacci.

Problema 1 Escriba los primeros cinco términos de otras tres sucesiones tipo Fibonacci. Nuevamente por analogía con los vectores, defina la *suma* de dos sucesiones $\mathbf{x} = [x_0, x_1, x_2, \ldots)$ y $\mathbf{y} = [y_0, y_1, y_2, \ldots)$ como la sucesión

$$\mathbf{x} + \mathbf{y} = [x_0 + y_0, x_1 + y_1, x_2 + y_2, \ldots)$$

Si c es un escalar, del mismo modo se puede definir el múltiplo escalar de una sucesión mediante

$$c\mathbf{x} = [cx_0, cx_1, cx_2, \ldots)$$

Problema 2 (a) Con los ejemplos del problema 1 u otros, calcule las sumas de varios pares de sucesiones tipo Fibonacci. ¿Las sucesiones resultantes parecen ser del tipo Fibonacci?

(b) Calcule varios múltiplos escalares de sus sucesiones tipo Fibonacci del problema 1. ¿Las sucesiones resultantes parecen ser tipo Fibonacci?

Problema 3 (a) Demuestre que si $\mathbf{x}$ y $\mathbf{y}$ son sucesiones tipo Fibonacci, entonces también lo es $\mathbf{x} + \mathbf{y}$.

(b) Demuestre que si $\mathbf{x}$ es una sucesión tipo Fibonacci y c es un escalar, entonces $c\mathbf{x}$ también es una sucesión tipo Fibonacci.

Denote con *Fib* al conjunto de todas las sucesiones tipo Fibonnaci. El problema 3 muestra que, como $\mathbb{R}^n$, *Fib* es cerrado para la suma y multiplicación escalar. Los siguientes ejercicios muestran que *Fib* tiene mucho más en común con $\mathbb{R}^n$.

Problema 4 Revise las propiedades algebraicas de los vectores en el Teorema 1.1. ¿*Fib* satisface todas las propiedades? ¿Cuál sucesión tipo Fibonacci desempaña el papel de $\mathbf{0}$? Para una sucesión tipo Fibonacci $\mathbf{x}$, ¿cuál es $-\mathbf{x}$? ¿$-\mathbf{x}$ también es una sucesión tipo Fibonacci?

Problema 5 En $\mathbb{R}^n$ se tienen los vectores base estándar $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. La sucesión de Fibonacci $\mathbf{f} = [0, 1, 1, 2, \ldots)$ puede considerarse como el análogo de $\mathbf{e}_2$ porque sus primeros dos términos son 0 y 1. ¿Qué sucesión $\mathbf{e}$ en *Fib* tiene el papel de $\mathbf{e}_1$?

Problema 6 Sea $\mathbf{x} = [x_0, x_1, x_2, \ldots)$ una sucesión tipo Fibonacci. Demuestre que $\mathbf{x}$ es una combinación lineal de $\mathbf{e}$ y $\mathbf{f}$.

Problema 7 Demuestre que $\mathbf{e}$ y $\mathbf{f}$ son linealmente independientes. (Esto es, demuestre que, si $c\mathbf{e} + d\mathbf{f} = \mathbf{0}$, entonces $c = d = 0$.)

Problema 8 Dadas sus respuestas a los problemas 6 y 7, ¿cuál sería un valor sensible para asignar a la "dimensión" de *Fib*? ¿Por qué?

Problema 9 ¿Existen algunas sucesiones geométricas en *Fib*? Esto es, si

$$[1, r, r^2, r^3, \ldots)$$

es una sucesión tipo Fibonacci, ¿cuáles son los posibles valores de r?

Problema 10 Encuentre una "base" para *Fib* que consista de sucesiones geométricas tipo Fibonacci.

Problema 11 Con su respuesta al problema 10, ofrezca una deducción alternativa de la *fórmula de Binet* [fórmula (5) de la sección 4.6]:

$$f_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$



para los términos de la sucesión de Fibonacci $\mathbf{f} = [f_0, f_1, f_2, \ldots)$.[*Sugerencia:* exprese $\mathbf{f}$ en términos de la base del problema 10.]

La sucesión de Lucas recibe su nombre en honor de Edouard Lucas (vea la página 347).

La ***sucesión de Lucas*** es la sucesión tipo Fibonacci

$$\mathbf{l} = [l_0, l_1, l_2, l_3, \ldots) = [2, 1, 3, 4, \ldots)$$

Problema 12 Use la base del problema 10 para encontrar un análogo de la fórmula de Binet para el n-ésimo término l_n de la sucesión de Lucas.

Problema 13 Demuestre que las sucesiones de Fibonacci y de Lucas se relacionan mediante la identidad

$$f_{n-1} + f_{n+1} = l_n \text{ para } n \geq 1$$



[*Sugerencia:* las sucesiones tipo Fibonacci $\mathbf{f}^- = [1, 1, 2, 3, \ldots)$ y $\mathbf{f}^+ = [1, 0, 1, 1, \ldots)$ forman una base para *Fib*. (¿Por qué?)]

En esta introducción se vio que la colección *Fib* de todas las sucesiones tipo Fibonacci se comportan en muchos aspectos como $\mathbb{R}^2$, aun cuando los "vectores" en realidad son sucesiones infinitas. Esta útil analogía conduce a la noción general de un *espacio vectorial* que es el tema de este capítulo.

6.1 Espacios y subespacios vectoriales

En los capítulos 1 y 3 se vio que el álgebra de vectores y el álgebra de matrices son similares en muchos aspectos. En particular, es posible sumar tanto vectores como matrices, y ambos pueden multiplicarse por escalares. Las propiedades que resultan de estas dos operaciones (Teorema 1.1 y Teorema 3.2) son idénticas en ambos escenarios. En esta sección se usan dichas propiedades para definir "vectores" generalizados que surgen en una gran variedad de ejemplos. Al probar teoremas generales acerca de dichos "vectores", se probarán simultáneamente resultados acerca de todos esos ejemplos. Este es el verdadero poder del álgebra: su habilidad para tomar propiedades de un escenario concreto, como $\mathbb{R}^n$, y *abstraerlas* a un escenario general.

Definición Sea V un conjunto sobre el cual se definen dos operaciones, llamadas *suma* y *multiplicación por un escalar*. Si $\mathbf{u}$ y $\mathbf{v}$ están en V, la *suma* de $\mathbf{u}$ y $\mathbf{v}$ se denota mediante $\mathbf{u} + \mathbf{v}$, y si c es un escalar, el *múltiplo escalar* de $\mathbf{u}$ por c se denota mediante $c\mathbf{u}$. Si los siguientes axiomas se cumplen para todos $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ en V y para todos los escalares c y d, entonces V se llama ***espacio vectorial*** y sus elementos se llaman ***vectores***.

1. $\mathbf{u} + \mathbf{v}$ está en V. — Cerradura bajo la suma
2. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ — Conmutatividad
3. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ — Asociatividad
4. Existe un elemento $\mathbf{0}$ en V, llamado ***vector cero***, tal que $\mathbf{u} + \mathbf{0} = \mathbf{u}$.
5. Para cada $\mathbf{u}$ en V, existe un elemento $-\mathbf{u}$ en V tal que $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.
6. $c\mathbf{u}$ está en V. — Cerradura bajo multiplicación escalar
7. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ — Distributividad
8. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ — Distributividad
9. $c(d\mathbf{u}) = (cd)\mathbf{u}$
10. $1\mathbf{u} = \mathbf{u}$

Por lo general, al matemático alemán Hermann Grassmann (1809-1877) se le acredita la introducción de la idea de un espacio vectorial (aunque él no lo llamó de esa forma) en 1844. Por desgracia, su obra es muy difícil de leer y no recibe la atención que merece. Una persona que lo estudió fue el matemático italiano Giuseppe Peano (1858-1932). En su libro de 1888, *Calcolo Geometrico*, Peano clarificó la obra temprana de Grassmann y tendió los axiomas para un espacio vectorial como se le conoce en la actualidad. El libro de Peano también es notable por introducir las operaciones con conjuntos. Sus notaciones $\cup$, $\cap$ y $\in$ (para "unión", "intersección" y "es un elemento de") son las que se utilizan hoy día, aunque otros matemáticos no las aceptaron de inmediato. La definición axiomática de Peano de espacio vectorial también tuvo poca influencia durante muchos años. La aceptación llegó en 1918, después de que Hermann Weyl (1885-1955) lo repitiera en su libro *Space, Time, Matter*, una introducción a la teoría de la relatividad general de Einstein.

Comentarios

- Por "escalares" usualmente se entenderán los números reales. En concordancia, debe referirse a V como a un *espacio vectorial real* (o un *espacio vectorial sobre los números reales*). También es posible que los escalares sean números complejos o que pertenezcan a $\mathbb{Z}_p$, donde p es primo. En esos casos, V se conoce como *espacio vectorial complejo* o *espacio vectorial sobre* $\mathbb{Z}_p$, respectivamente. La mayoría de los ejemplos serán espacios vectoriales reales, de modo que por lo general se omitirá el adjetivo "real". Si a algo se le refiere como "espacio vectorial", suponga que trabajará en el sistema de números reales.

  De hecho, los escalares pueden elegirse de cualquier sistema numérico en el que, dicho burdamente, puede sumar, restar, multiplicar y dividir de acuerdo con las leyes usuales de la aritmética. En álgebra abstracta, a tal sistema numérico se le conoce como ***campo***.
- La definición de espacio vectorial no especifica en qué consiste el conjunto V. Tampoco especifica cómo son las operaciones llamadas "suma" y "multiplicación por un escalar". Con frecuencia, serán familiares, pero no necesitan serlo. Vea el ejemplo 6.6 y los ejercicios 5-7.

Ahora observará varios ejemplos de espacios vectoriales. En cada caso, es necesario especificar el conjunto V y las operaciones de suma y multiplicación por un escalar y verificar los axiomas 1 a 10. Es necesario poner particular atención a los axiomas 1 y 6 (ce-

rradura), axioma 4 (la existencia de un vector cero en V) y el axioma 5 (cada vector en V debe tener un negativo en V).

Ejemplo 6.1

Para cualquier $n \geq 1$, $\mathbb{R}^n$ es un espacio vectorial con las operaciones usuales de suma y multiplicación por un escalar. Los axiomas 1 y 6 son consecuencia de las definiciones de dichas operaciones y los axiomas restantes se deducen del Teorema 1.1.

Ejemplo 6.2

El conjunto de todas las matrices de 2×3 es un espacio vectorial con las operaciones usuales de suma matricial y multiplicación escalar matricial. Aquí los "vectores" en realidad son matrices. Se sabe que la suma de dos matrices de 2×3 también es una matriz de 2×3 y que al multiplicar una matriz de 2×3 por un escalar produce otra matriz de 2×3; por tanto, se tiene cerradura. Los axiomas restantes se deducen del Teorema 3.2. En particular, el vector cero $\mathbf{0}$ es la matriz cero de 2×3, y el negativo de una matriz A de 2×3 es justo la matriz $-A$ de 2×3.

No hay nada especial acerca de las matrices de 2×3. Para cualesquiera enteros positivos m y n, el conjunto de todas las matrices de $m \times n$ forma un espacio vectorial con las operaciones usuales de suma matricial y multiplicación escalar matricial. Este espacio vectorial se denota M_{mn}.

Ejemplo 6.3

Sea $\mathscr{P}_2$ el conjunto de todos los polinomios de grado 2 o menor con coeficientes reales. Defina la suma y multiplicación por un escalar en la forma usual. (Vea el Apéndice D.) Si

$$p(x) = a_0 + a_1x + a_2x^2 \quad \text{y} \quad q(x) = b_0 + b_1x + b_2x^2$$

están en $\mathscr{P}_2$, entonces

$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2$$

tiene cuando mucho grado 2 y por tanto está en $\mathscr{P}_2$. Si c es un escalar, entonces

$$cp(x) = ca_0 + ca_1x + ca_2x^2$$

también está en $\mathscr{P}_2$. Esto verifica los axiomas 1 y 6.

El vector cero $\mathbf{0}$ es el polinomio cero; esto es, el polinomio cuyas entradas son todas cero. El negativo de un polinomio $p(x) = a_0 + a_1x + a_2x^2$ es el polinomio $-p(x) = -a_0 - a_1x - a_2x^2$. Ahora es fácil verificar los axiomas restantes. Se comprobará el axioma 2 y se dejarán los otros para el ejercicio 12. Con los anteriores $p(x)$ y $q(x)$, se tiene

$$\begin{aligned} p(x) + q(x) &= (a_0 + a_1x + a_2x^2) + (b_0 + b_1x + b_2x^2) \\ &= (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 \\ &= (b_0 + a_0) + (b_1 + a_1)x + (b_2 + a_2)x^2 \\ &= (b_0 + b_1x + b_2x^2) + (a_0 + a_1x + a_2x^2) \\ &= q(x) + p(x) \end{aligned}$$

donde la tercera igualdad es consecuencia del hecho de que la suma de números reales es conmutativa.

En general, para cualquier $n \geq 0$ fijo, el conjunto $\mathscr{P}_n$ de todos los polinomios de grado menor o igual a n es un espacio vectorial, como lo es el conjunto $\mathscr{P}$ de *todos* los polinomios.

Ejemplo 6.4

Sea $\mathscr{F}$ el conjunto de todas las funciones con valor real definidas en la recta numérica. Si f y g son dos de tales funciones y c es un escalar, entonces $f + g$ y cf están definidos por

$$(f + g)(x) = f(x) + g(x) \quad \text{y} \quad (cf)(x) = cf(x)$$

En otras palabras, el *valor* de $f + g$ en x se obtiene al sumar los valores de f y g en x [figura 6.1(a)]. De igual modo, el valor de cf en x es justo el valor de f en x multiplicado por el escalar c [figura 6.1(b)]. El vector cero en $\mathcal{F}$ es la función constante f_0 que es idénticamente cero; esto es, $f_0(x) = 0$ para toda x. El negativo de una función f es la función $-f$ definida por $(-f)(x) = -f(x)$ [figura 6.1(c)].

Los Axiomas 1 y 6 obviamente son verdaderos. La verificación de los axiomas restantes se deja como ejercicio 13. Por tanto, $\mathscr{F}$ es un espacio vectorial.

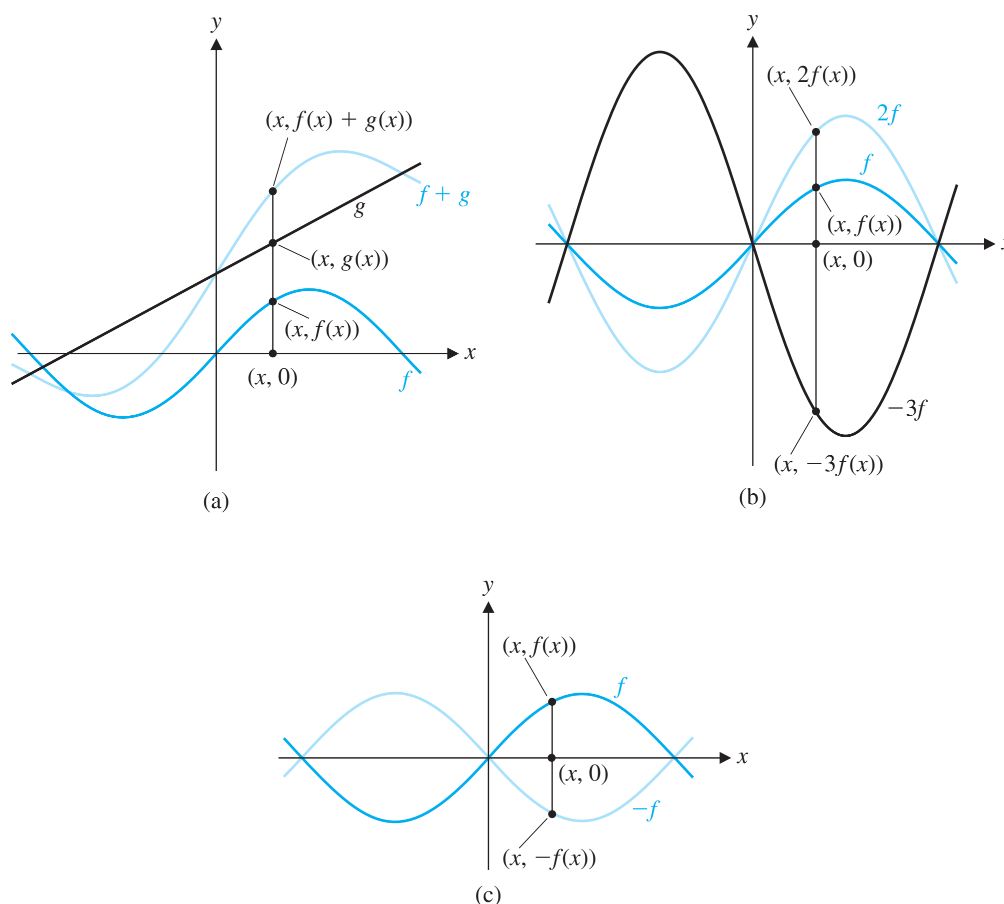


Figura 6.1
Las gráficas de (a) f, g y $f + g$, (b) f, $2f$ y $-3f$, y (c) f y $-f$

En el ejemplo 6.4 también podría considerar sólo aquellas funciones definidas en algún *intervalo cerrado* $[a, b]$ de la recta real. Este planteamiento también produce un espacio vectorial denotado mediante $\mathscr{F}[a, b]$.

Ejemplo 6.5

El conjunto $\mathbb{Z}$ de enteros con las operaciones usuales *no* es un espacio vectorial. Para demostrar esto, es suficiente con encontrar que falla *uno* de los diez axiomas y dar un ejemplo específico donde falle (un *contraejemplo*). En este caso, se descubre que no se tiene cerradura para la multiplicación por un escalar. Por ejemplo, el múltiplo del entero 2 por el escalar $\frac{1}{3}$ es $(\frac{1}{3})(2) = \frac{2}{3}$, que no es entero. Por tanto, no es cierto que cx esté en $\mathbb{Z}$ para *toda* x en $\mathbb{Z}$ y *todo* escalar c (esto es, falla el axioma 6).

Ejemplo 6.6

Sea $V = \mathbb{R}^2$ con la definición usual de suma, pero la siguiente definición de multiplicación por un escalar:

$$c\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ 0 \end{bmatrix}$$

Entonces, por ejemplo, $$1\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$



de modo que falla el axioma 10. [De hecho, los otros nueve axiomas son todos verdaderos (compruébelo), pero no es necesario sondear en ellos, porque V ya falló en ser un espacio vectorial. Este ejemplo demuestra el valor de anticiparse, en lugar de trabajar a lo largo de una lista de axiomas en el orden en el que se presentan.]

a + bi

Ejemplo 6.7

Sea $\mathbb{C}^2$ el conjunto de todos los pares ordenados de números complejos. Defina la suma y la multiplicación escalar como en $\mathbb{R}^2$, excepto que aquí los escalares son números complejos. Por ejemplo,

$$\begin{bmatrix} 1 + i \\ 2 - 3i \end{bmatrix} + \begin{bmatrix} -3 + 2i \\ 4 \end{bmatrix} = \begin{bmatrix} -2 + 3i \\ 6 - 3i \end{bmatrix}$$

y $$(1 - i)\begin{bmatrix} 1 + i \\ 2 - 3i \end{bmatrix} = \begin{bmatrix} (1 - i)(1 + i) \\ (1 - i)(2 - 3i) \end{bmatrix} = \begin{bmatrix} 2 \\ -1 - 5i \end{bmatrix}$$

Al usar las propiedades de los números complejos, es directo comprobar que se sostienen los diez axiomas. Por tanto, $\mathbb{C}^2$ es un espacio vectorial complejo.

En general, $\mathbb{C}^n$ es un espacio vectorial complejo para todo $n \geq 1$.

Ejemplo 6.8

Si p es primo, el conjunto $\mathbb{Z}_p^n$ (con las definiciones usuales de suma y multiplicación por escalares de $\mathbb{Z}_p$) es un espacio vectorial sobre $\mathbb{Z}_p$ para todo $n \geq 1$.

Antes de considerar más ejemplos, se enuncia un teorema que contiene algunas propiedades útiles de los espacios vectoriales. Es importante notar que al demostrar este teorema para espacios vectoriales en *general*, en realidad se demuestra para *todo* espacio vectorial *específico*.

Teorema 6.1 Sea V un espacio vectorial, $\mathbf{u}$ un vector en V y c un escalar.

a. $0\mathbf{u} = \mathbf{0}$
b. $c\mathbf{0} = \mathbf{0}$
c. $(-1)\mathbf{u} = -\mathbf{u}$
d. Si $c\mathbf{u} = \mathbf{0}$, entonces $c = 0$ o $\mathbf{u} = \mathbf{0}$.

Demostración Se demuestran las propiedades (b) y (d) y se dejan las demostraciones de las propiedades restantes como ejercicios.

(b) Se tiene

$$c\mathbf{0} = c(\mathbf{0} + \mathbf{0}) = c\mathbf{0} + c\mathbf{0}$$

por los axiomas 4 y 7 de espacio vectorial. Al sumar el negativo de $c\mathbf{0}$ a ambos lados se produce

$$c\mathbf{0} + (-c\mathbf{0}) = (c\mathbf{0} + c\mathbf{0}) + (-c\mathbf{0})$$

que implica

$$\begin{aligned} \mathbf{0} &= c\mathbf{0} + (c\mathbf{0} + (-c\mathbf{0})) && \text{por los axiomas 5 y 3} \\ &= c\mathbf{0} + \mathbf{0} && \text{por el axioma 5} \\ &= c\mathbf{0} && \text{por el axioma 4} \end{aligned}$$

(d) Suponga $c\mathbf{u} = \mathbf{0}$. Para demostrar que $c = 0$ o $\mathbf{u} = \mathbf{0}$, suponga que $c \neq 0$. (Si $c = 0$, no hay nada que probar.) Entonces, dado que $c \neq 0$, su recíproco $1/c$ está definido, y

$$\begin{aligned} \mathbf{u} &= 1\mathbf{u} && \text{por el axioma 10} \\ &= \left(\frac{1}{c}c\right)\mathbf{u} \\ &= \frac{1}{c}(c\mathbf{u}) && \text{por el axioma 9} \\ &= \frac{1}{c}\mathbf{0} \\ &= \mathbf{0} && \text{por la propiedad (b)} \end{aligned}$$

Se escribirá $\mathbf{u} - \mathbf{v}$ para $\mathbf{u} + (-\mathbf{v})$, lo que por tanto define la ***resta*** de vectores. También se explotará la propiedad de asociatividad de la suma para escribir sin ambigüedades $\mathbf{u} + \mathbf{v} + \mathbf{w}$ para la suma de tres vectores y, de manera más general,

$$c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n$$

para una ***combinación lineal*** de vectores.

Subespacios

Ya vio que, en $\mathbb{R}^n$, es posible que un espacio vectorial se asiente dentro de otro, lo que da lugar a la noción de subespacio. Por ejemplo, un plano que pasa por el origen es un subespacio de $\mathbb{R}^3$. Ahora se extiende este concepto a espacios vectoriales en general.

Definición Un subconjunto W de un espacio vectorial V se llama ***subespacio*** de V si W es en sí mismo un espacio vectorial con los mismos escalares, suma y multiplicación por un escalar que V.

Como en $\mathbb{R}^n$, comprobar si un subconjunto W de un espacio vectorial V es un subespacio de V implica probar sólo dos de los diez axiomas de espacio vectorial. Esta observación se demuestra como teorema.

Teorema 6.2 Sea V un espacio vectorial y sea W un subconjunto no vacío de V. Entonces W es un subespacio de V si y sólo si se cumplen las siguientes condiciones:

a. Si $\mathbf{u}$ y $\mathbf{v}$ están en W, entonces $\mathbf{u} + \mathbf{v}$ está en W.
b. Si $\mathbf{u}$ está en W y c es un escalar, entonces $c\mathbf{u}$ está en W.

Demostración Suponga que W es un subespacio de V. Entonces W satisface los axiomas 1 a 10 de espacio vectorial. En particular, el axioma 1 es la condición (a) y el axioma 6 es la condición (b).

Por el contrario, suponga que W es un subconjunto de un espacio vectorial V que satisface las condiciones (a) y (b). Por hipótesis, los axiomas 1 y 6 se cumplen. Los axiomas 2, 3, 7, 8, 9 y 10 se sustentan en W porque son verdaderos para *todos* los vectores en V y por tanto son verdaderos en particular para aquellos vectores en W. (Se dice que W *hereda* dichas propiedades de V.) Esto deja por probar los axiomas 4 y 5.

Dado que W no está vacío, contiene al menos un vector $\mathbf{u}$. Entonces la condición (b) y el Teorema 6.1(a) implican que $0\mathbf{u} = \mathbf{0}$ también está en W. Este es el axioma 4.

Si $\mathbf{u}$ está en V, entonces, al tomar $c = -1$ en la condición (b), se tiene que $-\mathbf{u} = (-1)\mathbf{u}$ también está en W, al usar el Teorema 6.1(c).

Comentario Dado que el Teorema 6.2 generaliza la noción de subespacio a partir del contexto de $\mathbb{R}^n$ a espacios vectoriales en general, todos los subespacios de $\mathbb{R}^n$ que se encontraron en el capítulo 3 son subespacios de $\mathbb{R}^n$ en el contexto actual. En particular, las rectas y planos que pasan por el origen son subespacios de $\mathbb{R}^3$.

Ejemplo 6.9 Ya se demostró que el conjunto $\mathscr{P}_n$ de todos los polinomios con grado n a lo más es un espacio vectorial. Por tanto, $\mathscr{P}_n$ es un subespacio del espacio vectorial $\mathscr{P}$ de *todos* los polinomios.

Ejemplo 6.10 Sea W en conjunto de matrices simétricas de $n \times n$. Demuestre que W es un subespacio de M_{nn}.

Solución Claramente, W no está vacío, así que sólo necesita demostrar las condiciones (a) y (b) en el Teorema 6.2. Sean A y B que están en W y sea c un escalar. Entonces $A^T = A$ y $B^T = B$, de donde se tiene que

$$(A + B)^T = A^T + B^T = A + B$$

Por tanto, $A + B$ es simétrica y, en consecuencia, está en W. De igual modo,

$$(cA)^T = cA^T = cA$$

de modo que cA es simétrica y, en consecuencia, está en W. Se demostró que W es cerrada para la suma y multiplicación por un escalar. Por tanto, es un subespacio de M_{nn}, por el Teorema 6.2.



Ejemplo 6.11

Sea $\mathscr{C}$ el conjunto de todas las funciones continuas con valor real definidas en $\mathbb{R}$ y sea $\mathscr{D}$ el conjunto de todas las funciones derivables con valor real en $\mathbb{R}$. Demuestre que $\mathscr{C}$ y $\mathscr{D}$ son subespacios de $\mathscr{F}$, el espacio vectorial de todas las funciones con valor real definidas en $\mathbb{R}$.

Solución Del cálculo se sabe que si f y g son funciones continuas y c es un escalar, entonces $f + g$ y cf también son continuas. Por tanto, $\mathscr{C}$ es cerrado bajo la suma y la multiplicación por un escalar y por tanto es un subespacio de $\mathscr{F}$. Si f y g son derivables, entonces también lo son $f + g$ y cf. De hecho,

$$(f + g)' = f' + g' \quad \text{y} \quad (cf)' = c(f')$$

De modo que $\mathscr{D}$ también es cerrado para la suma y la multiplicación por un escalar, lo que lo hace un subespacio de $\mathscr{F}$.

Es un teorema del cálculo que toda función derivable es continua. En consecuencia, $\mathscr{D}$ está contenido en $\mathscr{C}$ (denotado mediante $\mathscr{D} \subset \mathscr{C}$), lo que hace a $\mathscr{D}$ un subespacio de $\mathscr{C}$. También es el caso de que toda función polinominal es derivable, de modo que $\mathscr{P} \subset \mathscr{D}$, así que $\mathscr{P}$ es un subespacio de $\mathscr{D}$. Por tanto, se tiene una *jerarquía* de subespacios de $\mathscr{F}$, uno dentro del otro:

$$\mathscr{P} \subset \mathscr{D} \subset \mathscr{C} \subset \mathscr{F}$$

Esta jerarquía se muestra en la figura 6.2.

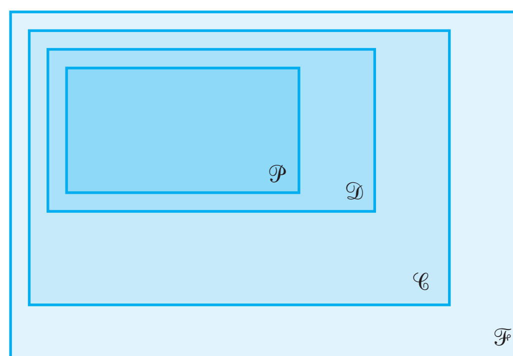
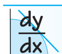


Figura 6.2
La jerarquía de subespacios de $\mathscr{F}$

Existen otros subespacios de $\mathscr{F}$ que pueden colocarse en esta jerarquía. Algunos de ellos se exploran en los ejercicios.

En la discusión anterior podría haberse restringido la atención a las funciones definidas en un intervalo cerrado $[a, b]$. Entonces los subespacios correspondientes de $\mathscr{F}[a, b]$ serían $\mathscr{C}[a, b]$, $\mathscr{D}[a, b]$ y $\mathscr{P}[a, b]$.



Ejemplo 6.12

Sea S el conjunto de todas las funciones que satisfacen la ecuación diferencial

$$f'' + f = 0 \tag{1}$$

Esto es, S es el conjunto solución de la ecuación (1). Demuestre que S es un subespacio de $\mathscr{F}$.

Solución S no está vacío, pues la función cero claramente satisface la ecuación (1). Sean f y g que están en S y sea c un escalar. Entonces

$$\begin{aligned}(f+g)'' + (f+g) &= (f'' + g'') + (f+g)\\ &= (f'' + f) + (g'' + g)\\ &= 0 + 0\\ &= 0\end{aligned}$$

lo que demuestra que $f + g$ está en S. De igual modo,

$$\begin{aligned}(cf)'' + cf &= cf'' + cf\\ &= c(f'' + f)\\ &= c0\\ &= 0\end{aligned}$$

de modo que cf también está en S.

Por tanto, S es cerrado para la suma y la multiplicación por un escalar y es un subespacio de $\mathscr{F}$.



La ecuación diferencial (1) es un ejemplo de una ***ecuación diferencial lineal homogénea***. Los conjuntos solución de tales ecuaciones siempre son subespacios de $\mathscr{F}$. Note que, en el ejemplo 6.12, en realidad no se resolvió la ecuación (1) (esto es, no se encontraron soluciones específicas distintas a la función cero). En la sección 6.7 se analizarán técnicas para encontrar soluciones para este tipo de ecuación.

Conforme obtiene experiencia al trabajar con espacios y subespacios vectoriales, notará que ciertos ejemplos tienden a parecerse entre sí. Por ejemplo, considere los espacios vectoriales $\mathbb{R}^4$, $\mathscr{P}_3$ y M_{22}. Los elementos típicos de estos espacios vectoriales son, respectivamente,

En palabras de Yogi Berra: "Nuevamente es déjà vu".

$$\mathbf{u} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \quad p(x) = a + bx + cx^2 + dx^3, \quad \text{y} \quad A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Cualquier cálculo que involucre las operaciones de espacio vectorial de suma y multiplicación por un escalar es en esencia el mismo en los tres escenarios. Para destacar las similitudes, en el siguiente ejemplo se realizarán los pasos necesarios en los tres espacios vectoriales lado por lado.

Ejemplo 6.13

(a) Demuestre que el conjunto W de todos los vectores de la forma

$$\begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix}$$

es un subespacio de $\mathbb{R}^4$.

(b) Demuestre que el conjunto W de todos los polinomios de la forma $a + bx - bx^2 + ax^3$ es un subespacio de $\mathscr{P}_3$.

(c) Demuestre que el conjunto W de todas las matrices de la forma $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ es un subespacio de M_{22}.

Solución

(a) W no es vacío porque contiene al vector cero $\mathbf{0}$. (Tome $a = b = 0$.) Sean $\mathbf{u}$ y $\mathbf{v}$ que están en W, por decir,

$$\mathbf{u} = \begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} c \\ d \\ -d \\ c \end{bmatrix}$$

Entonces

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} a + c \\ b + d \\ -b - d \\ a + c \end{bmatrix} = \begin{bmatrix} a + c \\ b + d \\ -(b + d) \\ a + c \end{bmatrix}$$

de modo que $\mathbf{u} + \mathbf{v}$ también está en W (porque tiene la *forma* correcta).

De igual modo, si k es un escalar, entonces

$$k\mathbf{u} = \begin{bmatrix} ka \\ kb \\ -kb \\ ka \end{bmatrix}$$

de modo que $k\mathbf{u}$ está en W.

Por tanto, W es un subconjunto no vacío de $\mathbb{R}^4$ que es cerrado bajo la suma y la multiplicación por un escalar. Por tanto, W es un subespacio de $\mathbb{R}^4$, por el Teorema 6.2.

(b) W no es vacío porque contiene el polinomio cero. (Tome $a = b = 0$.) Sean $p(x)$ y $q(x)$que están en W, por decir,

$$p(x) = a + bx - bx^2 + ax^3$$

y

$$q(x) = c + dx - dx^2 + cx^3$$

Entonces

$$p(x) + q(x) = (a + c) + (b + d)x - (b + d)x^2 + (a + c)x^3$$

de modo que $p(x) + q(x)$ también está en W (porque tiene la *forma* correcta).

De igual modo, si k es un escalar, entonces

$$kp(x) = ka + kbx - kbx^2 + kax^3$$

de modo que $kp(x)$ está en W.

Por tanto, W es un subconjunto no vacío de $\mathscr{P}_3$ que es cerrado para la suma y la multiplicación por un escalar. Por tanto, W es un subespacio de $\mathscr{P}_3$, por el Teorema 6.2.

(c) W no es vacío porque contiene la matriz O. (Tome $a = b = 0$.) Sean A y B en W, por decir,

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$$

y

$$B = \begin{bmatrix} c & d \\ -d & c \end{bmatrix}$$

Entonces

$$A + B = \begin{bmatrix} a + c & b + d \\ -(b + d) & a + c \end{bmatrix}$$

de modo que $A + B$ también está en W (porque tiene la *forma* correcta).

De igual modo, si k es un escalar, entonces

$$kA = \begin{bmatrix} ka & kb \\ -kb & ka \end{bmatrix}$$

de modo que kA está en W.

Por tanto, W es un subconjunto no vacío de M_{22} que es cerrado para la suma y la multiplicación por un escalar. Por tanto, W es un subespacio de M_{22}, por el Teorema 6.2.

El ejemplo 6.13 muestra que con frecuencia es posible relacionar ejemplos que, en la superficie, parecen no tener algo en común. En consecuencia, puede aplicar su conocimiento de $\mathbb{R}^n$ a polinomios, matrices y otros ejemplos. Esta idea se encontrará muchas veces en este capítulo y se precisará en la sección 6.5.

Ejemplo 6.14

Si V es un espacio vectorial, entonces V es claramente un subespacio de sí mismo. El conjunto $\{\mathbf{0}\}$, que consiste sólo del vector cero, también es un subespacio de V, llamado ***subespacio cero***. Para demostrar esto, note simplemente que se satisfacen las dos condiciones de cerradura del Teorema 6.2:

$$\mathbf{0} + \mathbf{0} = \mathbf{0} \quad \text{y} \quad c\mathbf{0} = \mathbf{0} \quad \text{para cualquier escalar } c$$

Los subespacios $\{\mathbf{0}\}$ y V se llaman ***subespacios triviales*** de V.

Un examen de la demostración del Teorema 6.2 revela el siguiente hecho útil:

Si W es un subespacio de un espacio vectorial V, entonces W contiene el vector cero $\mathbf{0}$ de V.

Este hecho es consistente con, y análogo a, el hecho de que las rectas y los planos son subespacios de $\mathbb{R}^3$ si y sólo si contienen al origen. El requisito de que todo subespacio debe contener $\mathbf{0}$ en ocasiones es útil para demostrar que un conjunto *no* es un subespacio.

Ejemplo 6.15 Sea W el conjunto de todas las matrices de 2×2 de la forma

$$\begin{bmatrix} a & a+1 \\ 0 & b \end{bmatrix}$$

¿W es un subespacio de M_{22}?

Solución Cada matriz en W tiene la propiedad de que su entrada $(1, 2)$ es uno más que su entrada $(1, 1)$. Dado que la matriz cero

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

no tiene esta propiedad, no está en W. Por tanto, W no es un subespacio de M_{22}.

Ejemplo 6.16 Sea W el conjunto de todas las matrices de 2×2 con determinante igual a 0. ¿W es un subespacio de M_{22}? (Dado que $O = 0$, la matriz cero está en W, de modo que el método del ejemplo 6.15 no es útil aquí.)

Solución Sea

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Entonces $\det A = \det B = 0$, de modo que A y B están en W. Pero

$$A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

de modo que $\det(A + B) = 1 \neq 0$, y por tanto $A + B$ no está en W. Por tanto, W no es cerrado para la suma y en consecuencia no es un subespacio de M_{22}.

Conjuntos generadores

La noción de conjunto generador de vectores se transporta fácilmente de $\mathbb{R}^n$ a espacios vectoriales en general.

Definición Si $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es un conjunto de vectores en un espacio vectorial V, entonces el conjunto de todas las combinaciones lineales de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ se llama el ***generador*** de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ y se denota $\text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ o $\text{gen}(S)$. Si $V = \text{gen}(S)$, entonces S se llama ***conjunto generador*** para V y se dice que V es ***generado*** por S.

Ejemplo 6.17

Demuestre que los polinomios 1, x y x^2 generan $\mathscr{P}_2$.

Solución Por su definición, un polinomio $p(x) = a + bx + cx^2$ es una combinación lineal de 1, x y x^2. Por tanto, $\mathscr{P}_2 = \text{gen}(1, x, x^2)$.

El ejemplo 6.17 puede generalizarse claramente para demostrar que $\mathscr{P}_n = \text{gen}(1, x, x^2, \ldots, x^n)$. Sin embargo, ningún conjunto finito de polinomios puede generar $\mathscr{P}$, el espacio vectorial de todos los polinomios. (Vea el ejercicio 44 de la sección 6.2.) Pero, si se permite que un conjunto generador sea infinito, entonces claramente el conjunto de *todas* las potencias no negativas de x lo hará. Esto es, $\mathscr{P} = \text{gen}(1, x, x^2, \ldots)$.

Ejemplo 6.18

Demuestre que $M_{23} = \text{gen}(E_{11}, E_{12}, E_{13}, E_{21}, E_{22}, E_{23})$, donde

$$E_{11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$E_{21} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad E_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad E_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(Esto es, E_{ij} es la matriz con un 1 en el renglón i, columna j y ceros en cualquiera otra parte.)

Solución Sólo es necesario observar que

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = a_{11}E_{11} + a_{12}E_{12} + a_{13}E_{13} + a_{21}E_{21} + a_{22}E_{22} + a_{23}E_{23}$$

Al extender este ejemplo, se ve que, en general, M_{mn} es generado por las mn matrices E_{ij}, donde $i = 1, \ldots, m$ y $j = 1, \ldots, n$.

Ejemplo 6.19

En $\mathscr{P}_2$, determine si $r(x) = 1 - 4x + 6x^2$ está en gen $(p(x), q(x))$, donde

$$p(x) = 1 - x + x^2 \quad \text{y} \quad q(x) = 2 + x - 3x^2$$

Solución Se buscan escalares c y d tales que $cp(x) + dq(x) = r(x)$. Esto significa que

$$c(1 - x + x^2) + d(2 + x - 3x^2) = 1 - 4x + 6x^2$$

Al reagrupar de acuerdo con las potencias de x, se tiene

$$(c + 2d) + (-c + d)x + (c - 3d)x^2 = 1 - 4x + 6x^2$$

Al igualar los coeficientes de potencias iguales de x se obtiene

$$\begin{aligned} c + 2d &= 1 \\ -c + d &= -4 \\ c - 3d &= 6 \end{aligned}$$

que se resuelve fácilmente para dar $c = 3$ y $d = -1$. Por tanto, $r(x) = 3p(x) - q(x)$, de modo que $r(x)$ está en gen$(p(x), q(x))$. (Compruébelo.)

Ejemplo 6.20 En $\mathscr{F}$, determine si sen $2x$ está en gen(sen x, cos x).

Solución Haga c sen $x + d$ cos $x =$ sen $2x$ y trate de determinar c y d de modo que esta ecuación sea verdadera. Puesto que se trata de funciones, la ecuación debe ser verdadera para *todos* los valores de x. Al hacer $x = 0$, se tiene

$$c \operatorname{sen} 0 + d \cos 0 = \operatorname{sen} 0 \quad \text{o} \quad c(0) + d(1) = 0$$

de donde se ve que $d = 0$. Al hacer $x = \pi/2$, se obtiene

$$c \operatorname{sen}(\pi/2) + d \cos(\pi/2) = \operatorname{sen}(\pi) \quad \text{o} \quad c(1) + d(0) = 0$$

lo que produce $c = 0$. Pero esto implica que sen $2x = 0$(sen x) $+ 0$(cos x) $= 0$ para toda x, lo que es absurdo, pues sen $2x$ no es la función cero. Se concluye que sen $2x$ no está en gen(sen x, cos x).

Comentario *Es* verdadero que sen $2x$ pueda escribirse en términos de sen x y cos x. Por ejemplo, se tiene la fórmula de doble ángulo sen $2x = 2$ sen x cos x. Sin embargo, ésta no es una combinación *lineal*.

Ejemplo 6.21 En M_{22}, describa el generador de $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ y $C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Solución Toda combinación lineal de A, B y C es de la forma

$$\begin{aligned} cA + dB + eC &= c\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + e\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} c+d & c+e \\ c+e & d \end{bmatrix} \end{aligned}$$

Esta matriz es simétrica, de modo que gen(A, B, C) está contenido dentro del subespacio de las matrices simétricas de 2×2. De hecho, se tiene una igualdad; esto es *toda* matriz simétrica de 2×2 está en gen(A, B, C). Para demostrar esto, sea $\begin{bmatrix} x & y \\ y & z \end{bmatrix}$ una matriz simétrica de 2×2. Al hacer

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} = \begin{bmatrix} c+d & c+e \\ c+e & d \end{bmatrix}$$

y resolver para c y d, se descubre que $c = x - z$, $d = z$ y $e = -x + y + z$. En consecuencia,

$$\begin{bmatrix} x & y \\ y & z \end{bmatrix} = (x - z)\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + z\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (-x + y + z)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

(Compruébelo.) Se concluye que gen(A, B, C) es el subespacio de las matrices simétricas de 2×2.

Como fue el caso en $\mathbb{R}^n$, el generador de un conjunto de vectores siempre es un subespacio del espacio vectorial que los contiene. El siguiente teorema precisa este resultado. Generaliza el Teorema 3.19.

Teorema 6.3 Sean $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ vectores en un espacio vectorial V.

a. $\text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ es un subespacio de V.
b. $\text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ es el subespacio más pequeño de V que contiene a $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$.

Demostración (a) La demostración de la propiedad (a) es idéntica a la del Teorema 3.19, con V en sustitución de $\mathbb{R}^n$.

(b) Para establecer la propiedad (b), es necesario demostrar que cualquier subespacio de V que contenga $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ también contiene $\text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$. En concordancia, sea W un subespacio de V que contiene $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Entonces, dado que W es cerrado para la suma y la multiplicación por un escalar, contiene a toda combinación lineal $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ de $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$. Por tanto, $\text{gen}(\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k)$ está contenido en W.

Ejercicios 6.1

En los ejercicios 1-11, determine si el conjunto dado, junto con las operaciones especificadas de suma y multiplicación por un escalar, es un espacio vectorial. Si no lo es, mencione todos los axiomas que no se cumplen.

1. El conjunto de todos los vectores en $\mathbb{R}^2$ de la forma $\begin{bmatrix} x \\ x \end{bmatrix}$, con la suma vectorial y la multiplicación por un escalar usuales.

2. El conjunto de todos los vectores $\begin{bmatrix} x \\ y \end{bmatrix}$ en $\mathbb{R}^2$ con $x \geq 0$, $y \geq 0$ (es decir el primer cuadrante), con la suma vectorial y la multiplicación por un escalar usuales.

3. El conjunto de todos los vectores $\begin{bmatrix} x \\ y \end{bmatrix}$ en $\mathbb{R}^2$ con $xy \geq 0$ (esto es la unión del primero y tercer cuadrantes), con la suma vectorial y la multiplicación por un escalar usuales.

4. El conjunto de todos los vectores $\begin{bmatrix} x \\ y \end{bmatrix}$ en $\mathbb{R}^2$ con $x \geq y$, con la suma vectorial y la multiplicación por un escalar usuales.

5. $\mathbb{R}^2$, con la suma usual pero con la multiplicación por un escalar definida por

$$c\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} cx \\ y \end{bmatrix}$$

6. $\mathbb{R}^2$, con la multiplicación por un escalar usual pero con suma definida por

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 + 1 \\ y_1 + y_2 + 1 \end{bmatrix}$$

7. El conjunto de todos los posibles números reales positivos, con suma $\oplus$ definida por $x \oplus y = xy$ y multiplicación por un escalar $\odot$ definida por $c \odot x = x^c$.

8. El conjunto de todos los números racionales, con la suma y multiplicación usuales.

9. El conjunto de todas las matrices triangulares superiores de 2×2, con la suma matricial y la multiplicación por un escalar usuales.

10. El conjunto de todas las matrices de 2×2 de la forma $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$, donde $ad = 0$, con la suma matricial y multiplicación por un escalar usuales.

11. El conjunto de todas las matrices antisimétricas de $n \times n$, con la suma matricial y la multiplicación por un escalar usuales (vea la página 168).

12. Termine de verificar que $\mathscr{P}_2$ es un espacio vectorial (vea el ejemplo 6.3).

13. Termine de verificar que $\mathscr{F}$ es un espacio vectorial (vea el ejemplo 6.4).

a + bi *En los ejercicios 14-17, determine si el conjunto dado, junto con las operaciones especificadas de suma y multiplicación por un escalar, es un espacio vectorial complejo. Si no lo es, mencione todos los axiomas que no se cumplen.*

14. El conjunto de todos los vectores en $\mathbb{C}^2$ de la forma $\begin{bmatrix} z \\ \bar{z} \end{bmatrix}$, con la suma vectorial y la multiplicación por un escalar usuales.

15. El conjunto $M_{mn}(\mathbb{C})$ de todas las matrices complejas de $m \times n$, con la suma matricial y la multiplicación por un escalar usuales.

16. El conjunto $\mathbb{C}^2$ con la suma usual pero la multiplicación por un escalar definida por $c\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} \bar{c}z_1 \\ \bar{c}z_2 \end{bmatrix}$.

17. $\mathbb{R}^n$, con la suma vectorial y multiplicación por un escalar usuales.

En los ejercicios 18-21, determine si el conjunto dado, junto con las operaciones especificadas de suma y multiplicación por un escalar, es un espacio vectorial sobre el $\mathbb{Z}_p$ indicado. Si no lo es, mencione todos los axiomas que no se sostienen.

18. El conjunto de todos los vectores en $\mathbb{Z}_2^n$ con un número par de números 1 sobre $\mathbb{Z}_2$ con la suma vectorial y la multiplicación por un escalar usuales.

19. El conjunto de todos los vectores en $\mathbb{Z}_2^n$ con número impar de números 1 sobre $\mathbb{Z}_2$ con la suma vectorial y la multiplicación por un escalar usuales.

20. El conjunto $M_{mn}(\mathbb{Z}_p)$ de todas las matrices de $m \times n$ con entradas de $\mathbb{Z}_p$, sobre $\mathbb{Z}_p$ con la suma matricial y la multiplicación por un escalar usuales.

21. $\mathbb{Z}_6$, sobre $\mathbb{Z}_3$ con la suma y la multiplicación usuales. (¡Piense esto cuidadosamente!)

22. Demuestre el Teorema 6.1(a).

23. Demuestre el Teorema 6.1(c).

En los ejercicios 24-45, use el Teorema 6.2 para determinar si W es un subespacio de V.

24. $V = \mathbb{R}^3$, $W = \left\{\begin{bmatrix} a \\ 0 \\ a \end{bmatrix}\right\}$

25. $V = \mathbb{R}^3$, $W = \left\{\begin{bmatrix} a \\ -a \\ 2a \end{bmatrix}\right\}$

26. $V = \mathbb{R}^3$, $W = \left\{\begin{bmatrix} a \\ b \\ a+b+1 \end{bmatrix}\right\}$

27. $V = \mathbb{R}^3$, $W = \left\{\begin{bmatrix} a \\ b \\ |a| \end{bmatrix}\right\}$

28. $V = M_{22}$, $W = \left\{\begin{bmatrix} a & b \\ b & 2a \end{bmatrix}\right\}$

29. $V = M_{22}$, $W = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} : ad \geq bc\right\}$

30. $V = M_{nn}$, $W = \{A \text{ en } M_{nn} : \det A = 1\}$

31. $V = M_{nn}$, W es el conjunto de matrices diagonales de $n \times n$

32. $V = M_{nn}$, W es el conjunto de matrices idempotentes de $n \times n$

33. $V = M_{nn}$, $W = \{A \text{ en } M_{nn} : AB = BA\}$, donde B es una matriz dada (fija)

34. $V = \mathcal{P}_2$, $W = \{bx + cx^2\}$

35. $V = \mathcal{P}_2$, $W = \{a + bx + cx^2 : a + b + c = 0\}$

36. $V = \mathcal{P}_2$, $W = \{a + bx + cx^2 : abc = 0\}$

37. $V = \mathcal{P}$, W es el conjunto de todos los polinomios de grado 3

38. $V = \mathcal{F}$, $W = \{f \text{ en } \mathcal{F} : f(-x) = f(x)\}$

39. $V = \mathcal{F}$, $W = \{f \text{ en } \mathcal{F} : f(-x) = -f(x)\}$

40. $V = \mathcal{F}$, $W = \{f \text{ en } \mathcal{F} : f(0) = 1\}$

41. $V = \mathcal{F}$, $W = \{f \text{ en } \mathcal{F} : f(0) = 0\}$

dy/dx **42.** $V = \mathcal{F}$, W es el conjunto de todas las funciones integrables

dy/dx **43.** $V = \mathcal{D}$, $W = \{f \text{ en } \mathcal{D} : f'(x) \geq 0 \text{ para toda } x\}$

dy/dx **44.** $V = \mathcal{F}$, $W = \mathcal{C}^{(2)}$, el conjunto de todas las funciones con segundas derivadas continuas

dy/dx **45.** $V = \mathcal{F}$, $W = \{f \text{ en } \mathcal{F} : \lim_{x \to 0} f(x) = \infty\}$

46. Sea V un espacio vectorial con subespacios U y W. Pruebe que $U \cap W$ es un subespacio de V.

47. Sea V un espacio vectorial con subespacios U y W. Ofrezca un ejemplo con $V = \mathbb{R}^2$ para demostrar que $U \cup W$ no necesita ser un subespacio de V.

48. Sea V un espacio vectorial con subespacios U y W. Defina la ***suma de U y W*** como

$$U + W = \{\mathbf{u} + \mathbf{w} : \mathbf{u} \text{ está en } U, \mathbf{w} \text{ está en } W\}$$

(a) Si $V = \mathbb{R}^3$, U es el eje x, y W es el eje y, ¿qué es $U + W$?

(b) Si U y W son subespacios de un espacio vectorial V, pruebe que $U + W$ es un subespacio de V.

49. Si U y V son espacios vectoriales, defina el ***producto cartesiano*** de U y V como

$$U \times V = \{(\mathbf{u}, \mathbf{v}) : \mathbf{u} \text{ está en } U \text{ y } \mathbf{v} \text{ está en } V\}$$

Demuestre que $U \times V$ es un espacio vectorial.

50. Sea W un subespacio de un espacio vectorial V. Pruebe que $\Delta = \{(\mathbf{w}, \mathbf{w}) : \mathbf{w}$ está en $W\}$ es un subespacio de $V \times V$.

En los ejercicios 51 y 52, sean $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$ *y* $B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$. *Determine si C está en gen(A, B).*

51. $C = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ **52.** $C = \begin{bmatrix} 3 & -5 \\ 5 & -1 \end{bmatrix}$

En los ejercicios 53 y 54, sean $p(x) = 1 - 2x$, $q(x) = x - x^2$ *y* $r(x) = -2 + 3x + x^2$. *Determine si* $s(x)$ *está en gen*$(p(x), q(x), r(x))$.

53. $s(x) = 3 - 5x - x^2$ **54.** $s(x) = 1 + x + x^2$

En los ejercicios 55-58, sean $f(x) = \text{sen}^2 x$ *y* $g(x) = \cos^2 x$. *Determine si* $h(x)$ *está en gen*$(f(x), g(x))$.

55. $h(x) = 1$ **56.** $h(x) = \cos 2x$

57. $h(x) = \text{sen}\, 2x$ **58.** $h(x) = \text{sen}\, x$

59. ¿ M_{22} es generado por $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$?

60. ¿ M_{22} es generado por $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$?

61. ¿ $\mathscr{P}_2$ es generado por $1 + x, x + x^2, 1 + x^2$?

62. ¿$\mathscr{P}_2$ es generado por $1 + x + 2x^2, 2 + x + 2x^2, -1 + x + 2x^2$?

63. Demuestre que todo espacio vectorial tiene un vector cero único.

64. Demuestre que para todo vector $\mathbf{v}$ en un espacio vectorial V, existe un $\mathbf{v}'$ único en V tal que $\mathbf{v} + \mathbf{v}' = \mathbf{0}$.

6.2 Independencia lineal, bases y dimensión

En esta sección se amplían las nociones de independencia lineal, base y dimensión a espacios vectoriales en general, lo que generaliza los resultados de las secciones 2.3 y 3.5. En la mayoría de los casos se realizan las demostraciones de los teoremas; simplemente se sustituye $\mathbb{R}^n$ por el espacio vectorial V.

Independencia lineal

Definición Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ en un espacio vectorial V es ***linealmente dependiente*** si existen escalares $c_1, c_2, \ldots, c_k$, *al menos uno de los cuales no es cero*, tales que

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

Un conjunto de vectores que no es linealmente dependiente se dice que es ***linealmente independiente***.

Como en $\mathbb{R}^n$, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ es linealmente independiente en un espacio vectorial V si y sólo si

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k = \mathbf{0} \quad \text{implica} \quad c_1 = 0, c_2 = 0, \ldots, c_k = 0$$

También se tiene la siguiente formulación alternativa de dependencia lineal que es muy útil.

Teorema 6.4 Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ en un espacio vectorial V es linealmente dependiente si y sólo si al menos uno de los vectores puede expresarse como una combinación lineal de los otros.

Demostración La demostración es idéntica a la del Teorema 2.5.

Como caso especial del Teorema 6.4, note que un conjunto de *dos* vectores es linealmente dependiente si y sólo si uno es un múltiplo escalar del otro.

Ejemplo 6.22 En $\mathscr{P}_2$, el conjunto $\{1 + x + x^2, 1 - x + 3x^2, 1 + 3x - x^2\}$ es linealmente dependiente, pues

$$2(1 + x + x^2) - (1 - x + 3x^2) = 1 + 3x - x^2$$

Ejemplo 6.23 En M_{22}, sea

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$$

Entonces $A + B = C$, de modo que el conjunto $\{A, B, C\}$ es linealmente dependiente.

Ejemplo 6.24 En $\mathscr{F}$, el conjunto $\{\operatorname{sen}^2 x, \cos^2 x, \cos 2x\}$ es linealmente dependiente, pues

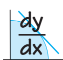
$$\cos 2x = \cos^2 x - \operatorname{sen}^2 x$$

Ejemplo 6.25 Demuestre que el conjunto $\{1, x, x^2, \ldots, x^n\}$ es linealmente independiente en $\mathscr{P}_n$.

Solución 1 Suponga que $c_0, c_1, \ldots, c_n$ son escalares tales que

$$c_0 \cdot 1 + c_1 x + c_2 x^2 + \cdots + c_n x^n = 0$$

Entonces el polinomio $p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$ es cero para todos los valores de x. Pero un polinomio de grado n como máximo no puede tener más de n ceros (vea el Apéndice D). De modo que $p(x)$ debe ser el polinomio cero, lo que significa que $c_0 = c_1 = c_2 = \cdots = c_n = 0$. Por tanto, $\{1, x, x^2, \ldots, x^n\}$ es linealmente independiente.



Solución 2 Comience, como en la primera solución, por suponer que

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n = 0$$

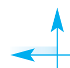
Dado que esto es verdadero para toda x, puede sustituir $x = 0$ para obtener $c_0 = 0$. Esto deja

$$c_1 x + c_2 x^2 + \cdots + c_n x^n = 0$$

Al sacar derivadas se obtiene

$$c_1 + 2c_2x + 3c_3x^2 + \cdots + nc_nx^{n-1} = 0$$

y al hacer $x = 0$, se ve que $c_1 = 0$. Al derivar $2c_2x + 3c_3x^2 + \cdots + nc_nx^{n-1} = 0$ y hacer $x = 0$, se encuentra que $2c_2 = 0$, de modo que $c_2 = 0$. Al continuar de esta forma, se encuentra que $k!c_k = 0$ para $k = 0, \ldots, n$. En consecuencia, $c_0 = c_1 = c_2 = \cdots = c_n = 0$ y $\{1, x, x^2, \ldots, x^n\}$ es linealmente independiente.



Ejemplo 6.26 En $\mathscr{P}_2$, determine si el conjunto $\{1 + x, x + x^2, 1 + x^2\}$ es linealmente independiente.

Solución Sean c_1, c_2 y c_3 escalares tales que

$$c_1(1 + x) + c_2(x + x^2) + c_3(1 + x^2) = 0$$

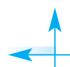
Entonces

$$(c_1 + c_3) + (c_1 + c_2)x + (c_2 + c_3)x^2 = 0$$

Esto implica que

$$\begin{aligned} c_1 + \quad\;\; c_3 &= 0 \\ c_1 + c_2 \quad\;\; &= 0 \\ c_2 + c_3 &= 0 \end{aligned}$$

cuya solución es $c_1 = c_2 = c_3 = 0$. Se tiene que $\{1 + x, x + x^2, 1 + x^2\}$ es linealmente independiente.



Comentario Compare el ejemplo 6.26 con el ejemplo 2.23(b). El sistema de ecuaciones que surge es exactamente el mismo. Esto se debe a la correspondencia entre $\mathscr{P}_2$ y $\mathbb{R}^3$ que relaciona

$$1 + x \leftrightarrow \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad x + x^2 \leftrightarrow \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad 1 + x^2 \leftrightarrow \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

y produce las columnas de la matriz de coeficientes del sistema lineal que debe resolverse. Por tanto, demostrar que $\{1 + x, x + x^2, 1 + x^2\}$ es linealmente independiente equivale a demostrar que

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

es linealmente independiente. Esto puede hacerse al simplemente establecer que la matriz

$$\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

tiene rank 3, por el teorema fundamental de las matrices invertibles.

Ejemplo 6.27

En $\mathscr{F}$ determine si el conjunto $\{\text{sen } x, \cos x\}$ es linealmente independiente.

Solución Las funciones $f(x) = \text{sen } x$ y $g(x) = \cos x$ son linealmente *dependientes* si y sólo si una de ellas es un múltiplo escalar de la otra. Pero es claro a partir de sus gráficas que este no es el caso, pues, por ejemplo, cualquier múltiplo distinto de cero de $f(x) = \text{sen } x$ tiene los mismos ceros, ninguno de los cuales es cero de $g(x) = \cos x$.

Es posible que el uso de este planteamiento no siempre sea adecuado, de modo que se ofrece el siguiente método directo más computacional. Suponga que c y d son escalares tales que

$$c \text{ sen } x + d \cos x = 0$$

Al hacer $x = 0$, se obtiene $d = 0$, y al hacer $x = \pi/2$, se obtiene $c = 0$. En consecuencia, el conjunto $\{\text{sen } x, \cos x\}$ es linealmente independiente.

Aunque las definiciones de dependencia e independencia lineal se enuncian en términos de conjuntos de vectores *finitos*, puede extender los conceptos a conjuntos *infinitos* del modo siguiente:

Un conjunto S de vectores en un espacio vectorial V es ***linealmente dependiente*** si contiene infinitos vectores linealmente dependientes. Un conjunto de vectores que no es linealmente dependiente se dice que es ***linealmente independiente***.

Note que para conjuntos finitos de vectores, esto es justo la definición original. El siguiente es un ejemplo de un conjunto infinito de vectores linealmente independientes.

Ejemplo 6.28

En $\mathscr{P}$, demuestre que $S = \{1, x, x^2, \ldots\}$ es linealmente independiente.

Solución Suponga que existe un subconjunto finito T de S que es linealmente dependiente. Sea x^m la potencia más alta de x en T y sea x^n la potencia más baja de x en T. Entonces existen escalares $c_n, c_{n+1}, \ldots, c_m$, no todos cero, tales que

$$c_n x^n + c_{n+1} x^{n+1} + \cdots + c_m x^m = 0$$

Pero, por un argumento similar al que se usó en el ejemplo 6.25, esto implica que $c_n = c_{n+1} = \cdots = c_m = 0$, lo que es una contradicción. Por tanto, S no puede contener finitos vectores linealmente dependientes, de modo que es linealmente independiente.

Bases

El importante concepto de base ahora se extiende con facilidad a espacios vectoriales arbitrarios.

Definición Un subconjunto $\mathcal{B}$ de un espacio vectorial V es una ***base*** para V si

1. $\mathcal{B}$ genera a V y
2. $\mathcal{B}$ es linealmente independiente.

Ejemplo 6.29

Si $\mathbf{e}_i$ es la i-ésima columna de la matriz identidad de $n \times n$, entonces $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ es una base para $\mathbb{R}^n$, llamada ***base estándar*** para $\mathbb{R}^n$.

Ejemplo 6.30

$\{1, x, x^2, \ldots, x^n\}$ es una base para $\mathscr{P}_n$ llamada ***base estándar*** para $\mathscr{P}_n$.

Ejemplo 6.31

El conjunto $\mathcal{E} = \{E_{11}, \ldots, E_{1n}, E_{21}, \ldots, E_{2n}, E_{m1}, \ldots, E_{mn}\}$ es una base para M_{mn}, donde las matrices E_{ij} se definen como en el ejemplo 6.18. $\mathcal{E}$ se llama ***base estándar*** para M_{mn}.

Ya se vio que $\mathcal{E}$ genera a M_{mn}. Es fácil demostrar que $\mathcal{E}$ es linealmente independiente. (¡Verifique esto!) Por tanto, $\mathcal{E}$ es una base para M_{mn}.

Ejemplo 6.32

Demuestre que $\mathcal{B} = \{1 + x, x + x^2, 1 + x^2\}$ es una base para $\mathscr{P}_2$.

Solución Ya se demostró en el ejemplo 6.26 que $\mathcal{B}$ es linealmente independiente. Para demostrar que $\mathcal{B}$ genera a $\mathscr{P}_2$, sea $a + bx + cx^2$ un polinomio arbitrario en $\mathscr{P}_2$. Debe demostrar que existen escalares c_1, c_2 y c_3 tales que

$$c_1(1 + x) + c_2(x + x^2) + c_3(1 + x^2) = a + bx + cx^2$$

o, de manera equivalente,

$$(c_1 + c_3) + (c_1 + c_2)x + (c_2 + c_3)x^2 = a + bx + cx^2$$

Al igualar coeficientes de potencias iguales de x, se obtiene el sistema lineal

$$\begin{aligned} c_1 + \phantom{c_2 + {}} c_3 &= a \\ c_1 + c_2 \phantom{{} + c_3} &= b \\ c_2 + c_3 &= c \end{aligned}$$

que tiene una solución, pues la matriz de coeficientes $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ tiene rank 3 y, por tanto, es invertible. (No es necesario saber *cuál* es la solución; sólo es necesario saber que existe.) En consecuencia, $\mathcal{B}$ es una base para $\mathscr{P}_2$.

Comentario Observe que la matriz $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ es la clave del ejemplo 6.32. Inmediatamente se le puede obtener al usar la correspondencia entre $\mathscr{P}_2$ y $\mathbb{R}^3$, como se indicó en el Comentario que siguió al ejemplo 6.26.

Ejemplo 6.33

Demuestre que $\mathcal{B} = \{1, x, x^2, \ldots\}$ es una base para $\mathcal{P}$.

Solución En el ejemplo 6.28 se vio que $\mathcal{B}$ es linealmente independiente. También genera a $\mathcal{P}$, pues claramente todo polinomio es una combinación lineal de (un número finito) de potencias de x.

Ejemplo 6.34

Encuentre bases para los tres espacios vectoriales del ejemplo 6.13:

(a) $W_1 = \left\{ \begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix} \right\}$ (b) $W_2 = \{a + bx - bx^2 + ax^3\}$ (c) $W_3 = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \right\}$

Solución Una vez más se trabajarán los tres ejemplos lado por lado para destacar las similitudes entre ellos. En un sentido estricto, se trata del *mismo* ejemplo, pero no será sino hasta la sección 6.5 cuando esta idea se precise a la perfección.

(a) Dado que

$$\begin{bmatrix} a \\ b \\ -b \\ a \end{bmatrix} = a\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

se tiene $W_1 = \text{gen}(\mathbf{u}, \mathbf{v})$, donde

$$\mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix}$$

Dado que $\{\mathbf{u}, \mathbf{v}\}$ es sin duda linealmente independiente, también es una base para W_1.

(b) Dado que

$$a + bx - bx^2 + ax^3 = a(1 + x^3) + b(x - x^2)$$

se tiene $W_2 = \text{gen}(u(x), v(x))$, donde

$$u(x) = 1 + x^3$$

y

$$v(x) = x - x^2$$

Dado que $\{u(x), v(x)\}$ sin duda es linealmente independiente, también es una base para W_2.

(c) Dado que

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

se tiene $W_3 = \text{gen}(U, V)$, donde

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{y} \quad V = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Dado que $\{U, V\}$ sin duda es linealmente independiente, también es una base para W_3.

Coordenadas

La sección 3.5 introdujo la idea de las coordenadas de un vector con respecto a una base para subespacios de $\mathbb{R}^n$. Ahora se extiende este concepto a espacios vectoriales arbitrarios.

Teorema 6.5

Sea V un espacio vectorial y sea $\mathcal{B}$ una base para V. Para todo vector $\mathbf{v}$ en V, existe exactamente una forma de escribir $\mathbf{v}$ como una combinación lineal de los vectores base en $\mathcal{B}$.

Demostración La demostración es igual que la del Teorema 3.29. Funciona incluso si la base $\mathcal{B}$ es infinita, pues las combinaciones lineales son, por definición, finitas.

El inverso del Teorema 6.5 también es verdadero. Esto es, si $\mathcal{B}$ es un conjunto de vectores en un espacio vectorial V con la propiedad de que todo vector en V se puede escribir de manera única como una combinación lineal de los vectores en $\mathcal{B}$, entonces $\mathcal{B}$ es una base para V (vea el ejercicio 30). En este sentido, la *propiedad de representación única* caracteriza a una base.

Dado que la representación de un vector con respecto a una base es única, la siguiente definición tiene sentido.

Definición Sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para un espacio vectorial V. Sea $\mathbf{v}$ un vector en V y escriba $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$. Entonces $c_1, c_2, \ldots, c_n$ se llaman las ***coordenadas de* v *con respecto a*** $\mathcal{B}$ y el vector columna

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

se llama ***vector coordenado de* v *con respecto a*** $\mathcal{B}$.

Observe que si la base $\mathcal{B}$ de V tiene n vectores, entonces $[\mathbf{v}]_{\mathcal{B}}$ es un vector (columna) en $\mathbb{R}^n$.

Ejemplo 6.35

Encuentre el vector coordenado $[p(x)]_{\mathcal{B}}$ de $p(x) = 2 - 3x + 5x^2$ con respecto a la base estándar $\mathcal{B} = \{1, x, x^2\}$ de $\mathcal{P}_2$.

Solución El polinomio $p(x)$ ya es una combinación lineal de 1, x y x^2, de modo que

$$[p(x)]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -3 \\ 5 \end{bmatrix}$$

La anterior es la correspondencia entre $\mathcal{P}_2$ y $\mathbb{R}^3$ que se destacó en un comentario después del ejemplo 6.26, y se puede generalizar fácilmente para demostrar que el vector coordenado de un polinomio

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \quad \text{en } \mathcal{P}_n$$

con respecto a la base estándar $\mathcal{B} = \{1, x, x^2, \ldots, x^n\}$ es justo el vector

$$[p(x)]_{\mathcal{B}} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{en } \mathbb{R}^{n+1}$$

Comentario El *orden* en el que los vectores base aparecen en $\mathcal{B}$ afecta el orden de las entradas o elementos en un vector coordenado. Por ejemplo, en el ejemplo 6.35, suponga

que los vectores base estándar están ordenados como $\mathcal{B}' = \{x^2, x, 1\}$. Entonces el vector coordenado de $p(x) = 2 - 3x + 5x^2$ con respecto a $\mathcal{B}'$ es

$$[p(x)]_{\mathcal{B}'} = \begin{bmatrix} 5 \\ -3 \\ 2 \end{bmatrix}$$

Ejemplo 6.36

Encuentre el vector coordenado $[A]_{\mathcal{B}}$ de $A = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix}$ con respecto a la base estándar $\mathcal{B} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ de M_{22}.

Solución Dado que

$$A = \begin{bmatrix} 2 & -1 \\ 4 & 3 \end{bmatrix} = 2\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + 4\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + 3\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$
$$= 2E_{11} - E_{12} + 4E_{21} + 3E_{22}$$

se tiene

$$[A]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -1 \\ 4 \\ 3 \end{bmatrix}$$

La anterior es la correspondencia entre M_{22} y $\mathbb{R}^4$ que se observó antes de la introducción al ejemplo 6.13. También se puede generalizar con facilidad para dar una correspondencia entre M_{mn} y $\mathbb{R}^{mn}$.

Ejemplo 6.37

Encuentre el vector coordenado $[p(x)]_{\mathcal{B}}$ de $p(x) = 1 + 2x - x^2$ con respecto a la base $\mathcal{C} = \{1 + x, x + x^2, 1 + x^2\}$ de $\mathscr{P}_2$.

Solución Necesita encontrar c_1, c_2 y c_3 tales que

$$c_1(1 + x) + c_2(x + x^2) + c_3(1 + x^2) = 1 + 2x - x^2$$

o, de manera equivalente,

$$(c_1 + c_3) + (c_1 + c_2)x + (c_2 + c_3)x^2 = 1 + 2x - x^2$$

Como en el ejemplo 6.32, esto significa que es necesario resolver el sistema

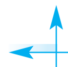
$$\begin{aligned} c_1 + c_3 &= 1 \\ c_1 + c_2 &= 2 \\ c_2 + c_3 &= -1 \end{aligned}$$

cuya solución se encontró era $c_1 = 2, c_2 = 0, c_3 = -1$. Por tanto,

$$[p(x)]_{\mathcal{C}} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$$



[Dado que este resultado dice que $p(x) = 2(1 + x) - (1 + x^2)$, es fácil comprobar que esto es correcto.]



El siguiente teorema muestra que el proceso de formar vectores coordenados es compatible con las operaciones de espacio vectorial de suma y multiplicación por un escalar.

Teorema 6.6 Sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para un espacio vectorial V. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en V y sea c un escalar. Entonces

a. $[\mathbf{u} + \mathbf{v}]_{\mathcal{B}} = [\mathbf{u}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}}$
b. $[c\mathbf{u}]_{\mathcal{B}} = c[\mathbf{u}]_{\mathcal{B}}$

Demostración Comience por escribir $\mathbf{u}$ y $\mathbf{v}$ en términos de vectores base, por decir, como

$$\mathbf{u} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \quad \text{y} \quad \mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n$$

Entonces, al usar las propiedades de espacio vectorial, se tiene

$$\mathbf{u} + \mathbf{v} = (c_1 + d_1)\mathbf{v}_1 + (c_2 + d_2)\mathbf{v}_2 + \cdots + (c_n + d_n)\mathbf{v}_n$$

y

$$c\mathbf{u} = (cc_1)\mathbf{v}_1 + (cc_2)\mathbf{v}_2 + \cdots + (cc_n)\mathbf{v}_n$$

de modo que

$$[\mathbf{u} + \mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} c_1 + d_1 \\ c_2 + d_2 \\ \vdots \\ c_n + d_n \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} = [\mathbf{u}]_{\mathcal{B}} + [\mathbf{v}]_{\mathcal{B}}$$

y

$$[c\mathbf{u}]_{\mathcal{B}} = \begin{bmatrix} cc_1 \\ cc_2 \\ \vdots \\ cc_n \end{bmatrix} = c\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = c[\mathbf{u}]_{\mathcal{B}}$$

Un sencillo corolario al Teorema 6.6 afirma que los vectores coordenados preservan las combinaciones lineales:

$$[c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k]_{\mathcal{B}} = c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{u}_k]_{\mathcal{B}} \qquad (1)$$

En el ejercicio 31 se le pide probar este corolario.

El aspecto más útil de los vectores coordenados es que permiten transferir información desde un espacio vectorial general a $\mathbb{R}^n$, donde tiene a su disposición las herramientas de los capítulos 1 a 3. En las secciones 6.3 y 6.6 se explorarán con más detalle estas ideas. Por el momento, se tiene el siguiente teorema útil.

Teorema 6.7 Sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para un espacio vectorial V y sean $\mathbf{u}_1, \ldots, \mathbf{u}_k$ vectores en V. Entonces $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es linealmente independiente en V si y sólo si $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_k]_{\mathcal{B}}\}$ es linealmente independiente en $\mathbb{R}^n$.

Demostración Suponga que $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es linealmente independiente en V y sea

$$c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{u}_k]_{\mathcal{B}} = \mathbf{0}$$

en $\mathbb{R}^n$. Pero entonces se tiene

$$[c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k]_{\mathcal{B}} = \mathbf{0}$$

al usar la ecuación (1), de modo que las coordenadas del vector $c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k$ con respecto a $\mathcal{B}$ son todas cero. Esto es,

$$c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 0\mathbf{v}_n = \mathbf{0}$$

La independencia lineal de $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ ahora fuerza que $c_1 = c_2 = \cdots = c_k = 0$, de modo que $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_k]_{\mathcal{B}}\}$ es linealmente independiente.

La implicación inversa, que usa ideas similares, se deja como ejercicio 32.

Observe que, en el caso especial donde $\mathbf{u}_i = \mathbf{v}_i$, se tiene

$$\mathbf{v}_i = 0 \cdot \mathbf{v}_1 + \cdots + 1 \cdot \mathbf{v}_i + \cdots + 0 \cdot \mathbf{v}_n$$

de modo que $[\mathbf{v}_i]_{\mathcal{B}} = \mathbf{e}_i$ y $\{[\mathbf{v}_1]_{\mathcal{B}}, \ldots, [\mathbf{v}_n]_{\mathcal{B}}\} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ es la base estándar en $\mathbb{R}^n$.

Dimensión

La definición de dimensión es la misma para un espacio vectorial que para un subespacio de $\mathbb{R}^n$: el número de vectores en una base para el espacio. Dado que un espacio vectorial puede tener más de una base, es necesario demostrar que esta definición tiene sentido; es decir: es necesario establecer que diferentes bases para el mismo espacio vectorial contienen el mismo número de vectores.

La parte (a) del siguiente teorema generaliza el Teorema 2.8.

Teorema 6.8 Sea $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ una base para un espacio vectorial V.

a. Cualquier conjunto con más de n vectores en V debe ser linealmente dependiente.
b. Cualquier conjunto con menos de n vectores en V no puede generar a V.

Demostración (a) Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ un conjunto de vectores en V, con $m > n$. Entonces $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_m]_{\mathcal{B}}\}$ es un conjunto con más de n vectores en $\mathbb{R}^n$ y, por tanto, es linealmente dependiente, por el Teorema 2.8. Esto significa que también $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ es linealmente dependiente, por el Teorema 6.7.

(b) Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ un conjunto de vectores en V, con $m < n$. Entonces $S = \{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_m]_{\mathcal{B}}\}$ es un conjunto con menos de n vectores en $\mathbb{R}^n$. Ahora, gen$(\mathbf{u}_1, \ldots, \mathbf{u}_m) = V$ si y sólo si gen$(S) = \mathbb{R}^n$ (vea el ejercicio 33). Pero gen(S) es justo el espacio columna de la matriz de $n \times m$

$$A = [[\mathbf{u}_1]_{\mathcal{B}} \ \cdots \ [\mathbf{u}_m]_{\mathcal{B}}]$$

de modo que $\dim(\text{gen}(S)) = \dim(\text{col}(A)) \le m < n$. Por tanto, S no puede generar a $\mathbb{R}^n$, así que $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ no genera a V.

Ahora se amplía el Teorema 3.23.

Teorema 6.9 **El teorema de la base**

Si un espacio vectorial V tiene una base con n vectores, entonces toda base para V tiene exactamente n vectores.

La demostración del Teorema 3.23 también funciona aquí, virtualmente palabra por palabra. Sin embargo, es más fácil usar el Teorema 6.8.

Demostración Sea $\mathcal{B}$ una base para V con n vectores y sea $\mathcal{B}'$ otra base para V con m vectores. Por el Teorema 6.8, $m \leq n$; de otro modo, $\mathcal{B}'$ sería linealmente dependiente.

Ahora use el Teorema 6.8 con los roles de $\mathcal{B}$ y $\mathcal{B}'$ intercambiados. Dado que $\mathcal{B}'$ es una base de V con m vectores, el Teorema 6.8 implica que cualquier conjunto con más de m vectores en V es linealmente dependiente. Por tanto, $n \leq m$, pues $\mathcal{B}$ es una base y, por tanto, es linealmente independiente.

Dado que $n \leq m$ y $m \leq n$ se debe tener $n = m$, como se requiere.

Ahora la siguiente definición tiene sentido, pues el número de vectores en una base (finita) no depende de la elección de base.

Definición Un espacio vectorial V se llama ***de dimensión finita*** si tiene una base que consiste de un número finito de vectores. La ***dimensión*** de V, denotada mediante dim V, es el número de vectores en una base para V. La dimensión del espacio vectorial cero $\{\mathbf{0}\}$ se define como cero. Un espacio vectorial que no tiene base finita se llama ***de dimensión infinita***.

Ejemplo 6.38

Dado que la base estándar para $\mathbb{R}^n$ tiene n vectores, dim $\mathbb{R}^n = n$. En el caso de $\mathbb{R}^3$, un subespacio unidimensional es justo el generador de un solo vector distinto de cero y por tanto es una recta que pasa por el origen. Un subespacio bidimensional es generado por su base de dos vectores linealmente independientes (es decir, no paralelos) y por tanto es un plano que pasa por el origen. Cualesquiera tres vectores linealmente independientes deben generar $\mathbb{R}^3$, por el teorema fundamental. Los subespacios de $\mathbb{R}^3$ ahora están completamente clasificados de acuerdo con la dimensión, como se muestra en la tabla 6.1.

Tabla 6.1

dim V	V
3	$\mathbb{R}^3$
2	Plano que pasa por el origen
1	Recta que pasa por el origen
0	$\{\mathbf{0}\}$

Ejemplo 6.39

La base estándar para $\mathscr{P}_n$ contiene $n + 1$ vectores (vea el ejemplo 6.30), de modo que dim $\mathscr{P}_n = n + 1$.

Ejemplo 6.40

La base estándar para M_{mn} contiene mn vectores (vea el ejemplo 6.31), de modo que $\dim M_{mn} = mn$.

Ejemplo 6.41

Tanto $\mathscr{P}$ como $\mathscr{F}$ son de dimensión infinita, pues cada uno de ellos contiene el conjunto infinito linealmente independiente $\{1, x, x^2, \ldots\}$ (vea el ejercicio 44).

Ejemplo 6.42

Encuentre la dimensión del espacio vectorial W de matrices simétricas de 2×2 (vea el ejemplo 6.10).

Solución Una matriz simétrica de 2×2 es de la forma

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

de modo que W es generado por el conjunto

$$S = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$$

Si S es linealmente independiente, entonces será una base para W. Al hacer

$$a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

se obtiene

$$\begin{bmatrix} a & b \\ b & c \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

de donde se tiene inmediatamente que $a = b = c = 0$. En consecuencia, S es linealmente independiente y, por tanto, es una base para W. Se concluye que $\dim W = 3$.

La dimensión de un espacio vectorial es su "número mágico". Conocer la dimensión de un espacio vectorial V brinda mucha información acerca de V y puede simplificar enormemente el trabajo necesario en ciertos tipos de cálculos, como ilustran los siguientes teoremas y ejemplos.

Teorema 6.10

Sea V un espacio vectorial con $\dim V = n$. Entonces:

a Cualquier conjunto linealmente independiente en V contiene cuando mucho n vectores.
b. Cualquier conjunto generador para V contiene al menos n vectores.
c. Cualquier conjunto linealmente independiente de exactamente n vectores en V es una base para V.
d. Cualquier conjunto generador para V que consiste de exactamente n vectores es una base para V.
e. Cualquier conjunto linealmente independiente en V se puede extender a una base para V.
f. Cualquier conjunto generador para V puede reducirse a una base para V.

Demostración Las demostraciones de las propiedades (a) y (b) se deducen de los incisos (a) y (b) del Teorema 6.8, respectivamente.

(c) Sea S un conjunto linealmente independiente de exactamente n vectores en V. Si S no genera a V, entonces existe algún vector $\mathbf{v}$ en V que no es una combinación lineal de los vectores en S. Al insertar $\mathbf{v}$ en S se produce un conjunto S' con $n + 1$ vectores que todavía es linealmente independiente (vea el ejercicio 54). Pero esto es imposible, por el Teorema 6.8(a). Se concluye que S debe a generar V y por tanto es una base para V.

(d) Sea S un conjunto generador para V que consiste de exactamente n vectores. Si S es linealmente dependiente, entonces algún vector $\mathbf{v}$ en S es una combinación lineal de los otros. Al quitar $\mathbf{v}$ el conjunto S' queda con $n - 1$ vectores que todavía generan V (vea el ejercicio 55). Pero esto es imposible, por el Teorema 6.8(b). Se concluye que S debe ser linealmente independiente y por tanto es una base para V.

(e) Sea S un conjunto linealmente independiente de vectores en V. Si S genera V, es una base para V y por tanto consiste de exactamente n vectores, por el teorema de la base. Si S no genera a V, entonces, como en la demostración de la propiedad (c), existe algún vector $\mathbf{v}$ en V que no es una combinación lineal de los vectores en S. Al insertar $\mathbf{v}$ en S se produce un conjunto S' que todavía es linealmente independiente. Si S' todavía no genera a V, puede repetir el proceso y expandirlo a un conjunto linealmente independiente más grande. Con el tiempo, este proceso debe detenerse, pues ningún conjunto linealmente independiente en V puede contener más de n vectores, por el Teorema 6.8(a). Cuando el proceso se detiene, se tiene un conjunto linealmente independiente S^* que contiene a S y también genera a V. Por tanto, S^* es una base para V que se extiende a S.

(f) En el ejercicio 56 se le pide probar esta propiedad.

El Teorema 6.10 debe verse, en parte, como un dispositivo para ahorrar trabajo. En muchas instancias puede reducir dramáticamente la cantidad de trabajo necesaria para comprobar que un conjunto de vectores es linealmente independiente, un conjunto generador o una base.

Ejemplo 6.43

En cada caso, determine si S es una base para V.

(a) $V = \mathscr{P}_2,\ S = \{1 + x, 2 - x + x^2, 3x - 2x^2, -1 + 3x + x^2\}$

(b) $V = M_{22},\ S = \left\{\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}\right\}$

(c) $V = \mathscr{P}_2,\ S = \{1 + x, x + x^2, 1 + x^2\}$

Solución (a) Dado que $\dim(\mathscr{P}_2) = 3$ y S contiene cuatro vectores, S es linealmente dependiente, por el Teorema 6.10(a). En consecuencia, S no es una base para $\mathscr{P}_2$.

(b) Dado que $\dim(M_{22}) = 4$ y S contiene tres vectores, S no puede generar M_{22}, por el Teorema 6.10(b). Por tanto, S no es una base para M_{22}.

(c) Dado que $\dim(\mathscr{P}_2) = 3$ y S contiene tres vectores, S será una base para $\mathscr{P}_2$ si es linealmente independiente o si genera a $\mathscr{P}_2$, por el Teorema 6.10(c) o (d). Es más fácil demostrar que S es linealmente independiente; esto se hizo en el ejemplo 6.26. Por tanto, S es una base para $\mathscr{P}_2$. (Este es el mismo problema del ejemplo 6.32, ¡pero vea cuánto más sencillo se vuelve al usar el Teorema 6.10!)

Ejemplo 6.44

Extienda $\{1 + x, 1 - x\}$ a una base para $\mathscr{P}_2$.

Solución Note primero que $\{1 + x, 1 - x\}$ es linealmente independiente. (¿Por qué?) Dado que $\dim(\mathscr{P}_2) = 3$, es necesario un tercer vector, uno que no sea linealmente de-

pendiente de los dos primeros. Podría proceder como en la demostración del Teorema 6.10(e) para encontrar tal vector usando ensayo y error. Sin embargo, es más fácil en la práctica proceder en forma diferente.

El conjunto dado de vectores se amplía al agregar *toda* la base estándar para $\mathscr{P}_2$. Esto produce

$$S = \{1 + x, 1 - x, 1, x, x^2\}$$

Ahora S es linealmente dependiente, por el Teorema 6.10(a), de modo que es necesario eliminar algunos vectores; en este caso, dos. ¿Cuáles? Use el Teorema 6.10(f) y comience con el primer vector que se agregó, 1. Dado que $1 = \frac{1}{2}(1 + x) + \frac{1}{2}(1 - x)$, el conjunto $\{1 + x, 1 - x, 1\}$ es linealmente dependiente, de modo que se elimina 1. De igual manera, $x = \frac{1}{2}(1 + x) - \frac{1}{2}(1 - x)$, de modo que $\{1 + x, 1 - x, x\}$ también es linealmente dependiente. Finalmente, compruebe que $\{1 + x, 1 - x, x^2\}$ es linealmente independiente. (¿Puede ver una forma rápida para decir esto?) En consecuencia, $\{1 + x, 1 - x, x^2\}$ es una base para $\mathscr{P}_2$ que extiende $\{1 + x, 1 - x\}$.

En el ejemplo 6.42, el espacio vectorial W de matrices simétricas de 2×2 es un subespacio del espacio vectorial M_{22} de todas las matrices de 2×2. Como se demostró, $W = 3 \leq 4 = \dim M_{22}$. Éste es un ejemplo de un resultado general, como muestra el teorema final de esta sección.

Teorema 6.11 Sea W un subespacio de un espacio vectorial de dimensión finita V. Entonces:

a. W es de dimensión finita y $\dim W \leq \dim V$.
b. $\dim W = \dim V$ si y sólo si $W = V$.

Demostración (a) Sea $\dim V = n$. Si $W = \{\mathbf{0}\}$, entonces $\dim(W) = 0 \leq n = \dim V$. Si W es distinta de cero, entonces cualquier base $\mathcal{B}$ para V (que contiene n vectores) ciertamente genera a W, pues W está contenido en V. Pero $\mathcal{B}$ puede reducirse a una base $\mathcal{B}'$ para W (que contiene cuando mucho n vectores), por el Teorema 6.10(f). Por tanto, W es de dimensión finita y $\dim(W) \leq n = \dim V$.

(b) Si $W = V$, entonces ciertamente $\dim W = \dim V$. Por otra parte, si $\dim W = \dim V = n$, entonces cualquier base $\mathcal{B}$ para W consiste de exactamente n vectores. Pero entonces estos son n vectores linealmente independientes en V y, en consecuencia, una base para V, por el Teorema 6.10(c). Por tanto, $V = \text{gen}(\mathcal{B}) = W$.

Ejercicios 6.2

En los ejercicios 1-4, pruebe los conjuntos de matrices para independencia lineal en M_{22}. Para las que sean linealmente dependientes, exprese una de las matrices como una combinación lineal de las otras.

1. $\left\{\begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}\right\}$

2. $\left\{\begin{bmatrix} 2 & 4 \\ -3 & 2 \end{bmatrix}, \begin{bmatrix} -1 & -2 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}\right\}$

3. $\left\{\begin{bmatrix} -1 & 1 \\ -2 & 2 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ -1 & 7 \end{bmatrix}\right\}$

4. $\left\{\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$

En los ejercicios 5-9, pruebe los conjuntos de polinomios para independencia lineal. Para los que sean linealmente dependientes, exprese uno de los polinomios como una combinación lineal de los otros.

5. $\{x, 1 + x\}$ en $\mathscr{P}_1$

6. $\{1 + x, 1 - x^2, 1 + x + x^2\}$ en $\mathscr{P}_2$

7. $\{x, 2x - x^2, 3x + 2x^2\}$ en $\mathscr{P}_2$

8. $\{2x, 1 - x^3, x^2 - x^3, 1 - 2x + x^2\}$ en $\mathcal{P}_3$

9. $\{1 - 2x, 3x + x^2 - x^3, 1 + x^2 + 2x^3, 3 + 2x + 3x^3\}$ en $\mathcal{P}_3$

En los ejercicios 10-14, pruebe los conjuntos de funciones para independencia lineal en $\mathcal{F}$. Para las que son linealmente dependientes, exprese una de las funciones como una combinación lineal de las otras.

10. $\{1, \operatorname{sen} x, \cos x\}$

11. $\{1, \operatorname{sen}^2 x, \cos^2 x\}$

12. $\{e^x, e^{-x}\}$

13. $\{1, \ln(2x), \ln(x^2)\}$

14. $\{\operatorname{sen} x, \operatorname{sen} 2x, \operatorname{sen} 3x\}$

15. Si f y g están en $\mathscr{C}^{(1)}$, el espacio vectorial de todas las funciones con derivadas continuas, entonces el determinante

$$W(x) = \begin{vmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{vmatrix}$$

se llama ***wronskiano*** de f y g [en honor del matemático polaco-francés Jósef Maria Hoëné-Wronski (1776-1853), quien trabajó en la teoría de determinantes y la filosofía de las matemáticas]. Demuestre que f y g son linealmente independientes si su wronskiano no es idénticamente cero (esto es, si existe alguna x tal que $W(x) \neq 0$).

16. En general, el wronskiano de $f_1, \ldots, f_n$ en $\mathscr{C}^{(n-1)}$ es el determinante

$$W(x) = \begin{vmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{vmatrix}$$

y $f_1, \ldots, f_n$ son linealmente independientes, siempre que $W(x)$ no sea idénticamente cero. Repita los ejercicios 10-14 usando la prueba del wronskiano.

17. Sea $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ un conjunto de vectores linealmente independientes en un espacio vectorial V.

(a) ¿$\{\mathbf{u} + \mathbf{v}, \mathbf{v} + \mathbf{w}, \mathbf{u} + \mathbf{w}\}$ es linealmente independiente? Demuestre que lo es u ofrezca un contraejemplo para demostrar que no lo es.

(b) ¿$\{\mathbf{u} - \mathbf{v}, \mathbf{v} - \mathbf{w}, \mathbf{u} - \mathbf{w}\}$ es linealmente independiente? Demuestre que lo es u ofrezca un contraejemplo para demostrar que no lo es.

En los ejercicios 18-25, determine si el conjunto $\mathcal{B}$ es una base para el espacio vectorial V.

18. $V = M_{22}$, $\mathcal{B} = \left\{\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right\}$

19. $V = M_{22}$, $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right\}$

20. $V = M_{22}$,
$\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}\right\}$

21. $V = M_{22}$,
$\mathcal{B} = \left\{\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ -3 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix}\right\}$

22. $V = \mathcal{P}_2$, $\mathcal{B} = \{x, 1 - x, 1 + x + x^2\}$

23. $V = \mathcal{P}_2$, $\mathcal{B} = \{1 - x, 1 - x^2, x - x^2\}$

24. $V = \mathcal{P}_2$, $\mathcal{B} = \{2 + 3x - x^2, 1 + 2x + 2x^2\}$

25. $V = \mathcal{P}_2$, $\mathcal{B} = \{1, 2 - x, 3 - x^2, x + 2x^2\}$

26. Encuentre el vector coordenado de $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ con respecto a la base $\mathcal{B} = \{E_{12}, E_{11}, E_{22}, E_{21}\}$ de M_{22}.

27. Encuentre el vector coordenado de $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ con respecto a la base $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$ de M_{22}.

28. Encuentre el vector coordenado de $p(x) = 2 - x + 3x^2$ con respecto a la base $\mathcal{B} = \{1 + x, 1 - x, x^2\}$ de $\mathcal{P}_2$.

29. Encuentre el vector coordenado de $p(x) = 2 - x + 3x^2$ con respecto a la base $\mathcal{B} = \{1, 1 + x, -1 + x^2\}$ de $\mathcal{P}_2$.

30. Sea $\mathcal{B}$ un conjunto de vectores en un espacio vectorial V con la propiedad de que todo vector en V puede escribirse de manera única como una combinación lineal de los vectores en $\mathcal{B}$. Pruebe que $\mathcal{B}$ es una base para V.

31. Sea $\mathcal{B}$ una base para un espacio vectorial V, sean $\mathbf{u}_1, \ldots, \mathbf{u}_k$ vectores en V, y sean $c_1, \ldots, c_k$ escalares. Demuestre que $[c_1\mathbf{u}_1 + \cdots + c_k\mathbf{u}_k]_{\mathcal{B}} = c_1[\mathbf{u}_1]_{\mathcal{B}} + \cdots + c_k[\mathbf{u}_k]_{\mathcal{B}}$.

32. Concluya la demostración del Teorema 6.7 al demostrar que si $\{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_k]_{\mathcal{B}}\}$ es linealmente independiente en $\mathbb{R}^n$, entonces $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es linealmente independiente en V.

33. Sea $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ un conjunto de vectores en un espacio vectorial n-dimensional V y sea $\mathcal{B}$ una base para V. Sea $S = \{[\mathbf{u}_1]_{\mathcal{B}}, \ldots, [\mathbf{u}_m]_{\mathcal{B}}\}$ el conjunto de vectores coordenados de $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ con respecto a $\mathcal{B}$. Demuestre que $\operatorname{gen}(\mathbf{u}_1, \ldots, \mathbf{u}_m) = V$ si y sólo si $\operatorname{gen}(S) = \mathbb{R}^n$.

En los ejercicios 34-39, encuentre la dimensión del espacio vectorial V y proponga una base para V.

34. $V = \{p(x) \text{ en } \mathcal{P}_2 : p(0) = 0\}$

35. $V = \{p(x) \text{ en } \mathcal{P}_2 : p(1) = 0\}$

36. $V = \{p(x) \text{ en } \mathcal{P}_2 : xp'(x) = p(x)\}$

37. $V = \{A \text{ en } M_{22} : A \text{ es triangular superior}\}$

38. $V = \{A \text{ en } M_{22} : A \text{ es antisimétrica}\}$

39. $V = \{A \text{ en } M_{22} : AB = BA\}$, donde $B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

40. Encuentre una fórmula para la dimensión del espacio vectorial de las matrices simétricas de $n \times n$.

41. Encuentre una fórmula para la dimensión del espacio vectorial de las matrices antisimétricas de $n \times n$.

42. Sean U y W subespacios de un espacio vectorial de dimensión finita V. Demuestre la ***identidad de Grassmann***:

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

[*Sugerencia:* el subespacio $U + W$ se definió en el ejercicio 48 de la sección 6.1. Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ una base para $U \cap W$. Extienda $\mathcal{B}$ a una base $\mathcal{C}$ de U y una base $\mathcal{D}$ de W. Demuestre que $\mathcal{C} \cup \mathcal{D}$ es una base para $U + W$.]

43. Sean U y V espacios vectoriales de dimensión finita.

(**a**) Encuentre una fórmula para $\dim(U \times V)$ en términos de $\dim U$ y $\dim V$. (Vea el ejercicio 49 de la sección 6.1.)

(**b**) Si W es un subespacio de V, demuestre que $\dim \Delta = \dim W$, donde $\Delta = \{(\mathbf{w}, \mathbf{w}) : \mathbf{w} \text{ está en } W\}$.

44. Demuestre que el espacio vectorial $\mathscr{P}$ es de dimensión infinita. [*Sugerencia:* suponga que tiene una base finita. Demuestre que existe algún polinomio que no es una combinación lineal de esta base.]

45. Extienda $\{1 + x, 1 + x + x^2\}$ a una base para $\mathscr{P}_2$.

46. Extienda $\left\{\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\right\}$ a una base para M_{22}.

47. Extienda $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right\}$ a una base para M_{22}.

48. Extienda $\left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right\}$ a una base para el espacio vectorial de las matrices simétricas de 2×2.

49. Encuentre una base para $\text{gen}(1, 1 + x, 2x)$ en $\mathscr{P}_1$.

50. Encuentre una base para $\text{gen}(1 - 2x, 2x - x^2, 1 - x^2, 1 + x^2)$ en $\mathscr{P}_2$.

51. Encuentre una base para $\text{gen}(1 - x, x - x^2, 1 - x^2, 1 - 2x + x^2)$ en $\mathscr{P}_2$.

52. Encuentre una base para $\text{gen}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\right)$ en M_{22}.

53. Encuentre una base para $\text{gen}(\text{sen}^2 x, \cos^2 x, \cos 2x)$ en $\mathscr{F}$.

54. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto linealmente independiente en un espacio vectorial V. Demuestre que si $\mathbf{v}$ es un vector en V que no está en $\text{gen}(S)$, entonces $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{v}\}$ todavía es linealmente independiente.

55. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto generador para un espacio vectorial V. Demuestre que si $\mathbf{v}_n$ está en $\text{gen}(\mathbf{v}_1, \ldots, \mathbf{v}_{n-1})$, entonces $S' = \{\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}\}$ todavía es un conjunto generador de V.

56. Demuestre el Teorema 6.10(f).

57. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base para un espacio vectorial V y sean $c_1, \ldots, c_n$ escalares distintos de cero. Pruebe que $\{c_1\mathbf{v}_1, \ldots, c_n\mathbf{v}_n\}$ también es una base para V.

58. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base para un espacio vectorial V. Pruebe que

$$\{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3, \ldots, \mathbf{v}_1 + \cdots + \mathbf{v}_n\}$$

también es una base para V.

Sean $a_0, a_1, \ldots, a_n$, $n + 1$ distintos números reales. Defina polinomios $p_0(x), p_1(x), \ldots, p_n(x)$ mediante

$$p_i(x) = \frac{(x - a_0)\cdots(x - a_{i-1})(x - a_{i+1})\cdots(x - a_n)}{(a_i - a_0)\cdots(a_i - a_{i-1})(a_i - a_{i+1})\cdots(a_i - a_n)}$$

*Los anteriores se llaman **polinomios de Lagrange** asociados con $a_0, a_1, \ldots, a_n$. [Joseph-Louis Lagrange (1736-1813) nació en Italia pero pasó la mayor parte de su vida en Alemania y Francia. Realizó importantes aportaciones a campos como la teoría de números, álgebra, astronomía, mecánica y el cálculo de variaciones. En 1773, Lagrange fue el primero en proporcionar la interpretación de volumen de un determinante (vea el capítulo 4).]*

59. (**a**) Calcule los polinomios de Lagrange asociados con $a_0 = 1, a_1 = 2, a_2 = 3$.

(**b**) Demuestre, en general, que

$$p_i(a_j) = \begin{cases} 0 & \text{si } i \neq j \\ 1 & \text{si } i = j \end{cases}$$

60. (**a**) Pruebe que el conjunto $\mathcal{B} = \{p_0(x), p_1(x), \ldots, p_n(x)\}$ de polinomios de Lagrange es linealmente independiente en $\mathscr{P}_n$. [*Sugerencia:* establezca $c_0p_0(x) + \cdots c_np_n(x) = 0$ y use el ejercicio 59(b).]

(**b**) Deduzca que $\mathcal{B}$ es una base para $\mathscr{P}_n$.

61. Si $q(x)$ es un polinomio arbitrario en $\mathscr{P}_n$, se deduce del ejercicio 60(b) que

$$q(x) = c_0p_0(x) + \cdots + c_np_n(x) \qquad (1)$$

para algunos escalares $c_0, \ldots, c_n$.

(**a**) Demuestre que para $i = 0, \ldots, n$, $c_i = q(a_i)$ y deduzca que $q(x) = q(a_0)p_0(x) + \cdots + q(a_n)p_n(x)$ es la representación única de $q(x)$ con respecto a la base $\mathcal{B}$.

(b) Demuestre que, para cualesquier $n + 1$ puntos $(a_0, c_0), (a_1, c_1), \ldots, (a_n, c_n)$ con distintos primeros componentes, la función $q(x)$ definida por la ecuación (1) es el polinomio de grado único cuando mucho n que pasa a través de todos los puntos. Esta fórmula se conoce como la ***fórmula de interpolación de Lagrange***. (Compare esta fórmula con el problema 19 de la Exploración: Aplicaciones geométricas de los determinantes, en el capítulo 4.)

(c) Use la fórmula de interpolación de Lagrange para encontrar el polinomio de grado cuando mucho 2 que pasa a través de los puntos

(i) $(1, 6), (2, -1)$, y $(3, -2)$

(ii) $(-1, 10), (0, 5)$, y $(3, 2)$

62. Use la fórmula de interpolación de Lagrange para demostrar que, si un polinomio en $\mathscr{P}_n$ tiene $n + 1$ ceros, entonces debe ser el polinomio cero.

63. Encuentre una fórmula para el número de matrices invertibles en $M_{nn}(\mathbb{Z}_p)$. [*Sugerencia:* esto es igual que determinar el número de bases diferentes para $\mathbb{Z}_p^n$. (¿Por qué?) Cuente el número de formas para construir una base para $\mathbb{Z}_p^n$, un vector a la vez.]

Exploración

Cuadrados mágicos

El grabado que se muestra en la página 479 es *Melancolía I* (1514), de Albrecht Dürer. Entre los muchos artefactos matemáticos de este grabado está el cuadro de números que cuelga en la pared de la esquina superior derecha. (Se muestra agrandado en el inserto ilustrado.) Tal arreglo de números se conoce como *cuadrado mágico*. Puede considerarse como una matriz de 4×4

$$\begin{bmatrix} 16 & 3 & 2 & 13 \\ 5 & 10 & 11 & 8 \\ 9 & 6 & 7 & 12 \\ 4 & 15 & 14 & 1 \end{bmatrix}$$

Observe que los números en cada renglón, en cada columna y en ambas diagonales tienen la misma suma: 34. Observe también que las entradas son los enteros 1, 2, . . ., 16. (Note que Dürer colocó el 15 y el 14 uno junto al otro en el último renglón, lo que da la fecha del grabado.) Estas observaciones condujeron a la siguiente definición.

Definición Una matriz M de $n \times n$ se conoce como ***cuadrado mágico*** si la suma de las entradas es la misma en cada renglón, cada columna y ambas diagonales. Esta suma común se llama el ***peso*** de M, que se denota $\mathrm{wt}(M)$. Si M es un cuadrado mágico de $n \times n$ que contiene cada una de las entradas $1, 2, \ldots, n^2$ exactamente una vez, entonces M se llama ***cuadrado mágico clásico***.

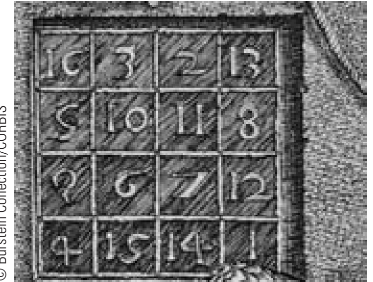
1. Si M es un cuadrado mágico clásico de $n \times n$, demuestre que

$$\mathrm{wt}(M) = \frac{n(n^2 + 1)}{2}$$

[*Sugerencia:* use el ejercicio 51 de la sección 2.4.]

2. Encuentre un cuadrado mágico clásico de 3×3. Encuentre uno diferente. ¿Sus dos ejemplos se relacionan en alguna forma?





3. Claramente, la matriz de 3×3 con todas las entradas iguales a $\frac{1}{3}$ es un cuadrado mágico con peso 1. Con su respuesta al problema 2, encuentre un cuadrado mágico de 3×3 con peso 1, *cuyas entradas sean todas diferentes*. Describa un método para construir un cuadrado mágico de 3×3 con distintas entradas y peso w para cualquier número real w.

Sea Mag_n el conjunto de todos los cuadrados mágicos de $n \times n$, y sea Mag_n^0 el conjunto de todos los cuadrados mágicos de $n \times n$ de peso 0.

4. (a) Demuestre que Mag_3 es un subespacio de M_{33}.
   (b) Demuestre que Mag_3^0 es un subespacio de Mag_3.

5. Use los problemas 3 y 4 para demostrar que si M es un cuadrado mágico de 3×3 con peso w, entonces puede escribir M como

$$M = M_0 + kJ$$

donde M_0 es un cuadrado mágico de 3×3 de peso 0, J es la matriz de 3×3 que consiste por completo de números 1 y k es un escalar. ¿Cuál debe ser k? [*Sugerencia:* demuestre que $M - kJ$ está en Mag_3^0 para un valor adecuado de k.]

Ahora se intentará encontrar una forma de describir *todos* los cuadrados mágicos de 3×3. Sea

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

un cuadrado mágico con peso 0. Las condiciones en los renglones, columnas y diagonales dan lugar a un sistema de ocho ecuaciones lineales homogéneas con las variables $a, b, \ldots, i$.

6. Escriba este sistema de ecuaciones y resuélvalo. [*Nota:* usar un CAS facilitará los cálculos.]

7. Encuentre la dimensión de Mag_3^0. *Sugerencia:* al hacer una sustitución, si es necesario, use su solución al problema 6 para demostrar que M puede escribirse en la forma

$$M = \begin{bmatrix} s & -s-t & t \\ -s+t & 0 & s-t \\ -t & s+t & -s \end{bmatrix}$$

8. Encuentre la dimensión de Mag_3. [*Sugerencia:* combine los resultados de los problemas 5 y 7.]

9. ¿Puede encontrar una forma directa de demostrar que la entrada (2, 2) de un cuadrado mágico de 3×3 con peso w debe ser $w/3$? [*Sugerencia:* sume y reste ciertos renglones, columnas y diagonales para dejar un múltiplo de la entrada central.]

10. Sea M un cuadrado mágico de 3×3 de peso 0, obtenido a partir de un cuadrado mágico clásico de 3×3, como en el problema 5. Si M tiene la forma dada en el problema 7, escriba una ecuación para la suma de los cuadrados de las entradas de M. Demuestre que esta es la ecuación de una circunferencia con las variables s y t, y grafíquela cuidadosamente. Demuestre que existen exactamente ocho puntos (s, t) en esta circunferencia con s y t enteros. Con el problema 8, demuestre que estos ocho puntos dan lugar a ocho cuadrados mágicos clásicos de 3×3. ¿Cómo se relacionan mutuamente estos cuadrados mágicos?

6.3 Cambio de base

En muchas aplicaciones, un problema descrito en un sistema coordenado puede resolverse más fácilmente al cambiar a un nuevo sistema coordenado. Este cambio por lo general se logra al realizar un cambio de variables, un proceso que probablemente usted encontró en otros cursos de matemáticas. En álgebra lineal, una base proporciona un sistema coordenado para un espacio vectorial, mediante la noción de vectores coordenados. La elección de la base correcta con frecuencia simplificará enormemente un problema particular. Por ejemplo, considere la estructura molecular del zinc, que se muestra en la figura 6.3(a). Un científico que estudia el zinc acaso quiera medir las longitudes de los enlaces entre los átomos, los ángulos entre dichos enlaces, etcétera. Tal análisis se facilitará enormemente al introducir coordenadas y utilizar las herramientas del álgebra lineal. La base estándar y los ejes coordenados estándar asociados *xyz* no siempre son la mejor opción. Como muestra la figura 6.3(b), en este caso probablemente $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ sea una mejor opción de base para $\mathbb{R}^3$ que la base estándar, pues dichos vectores se alinean muy bien con los enlaces entre los átomos de zinc.

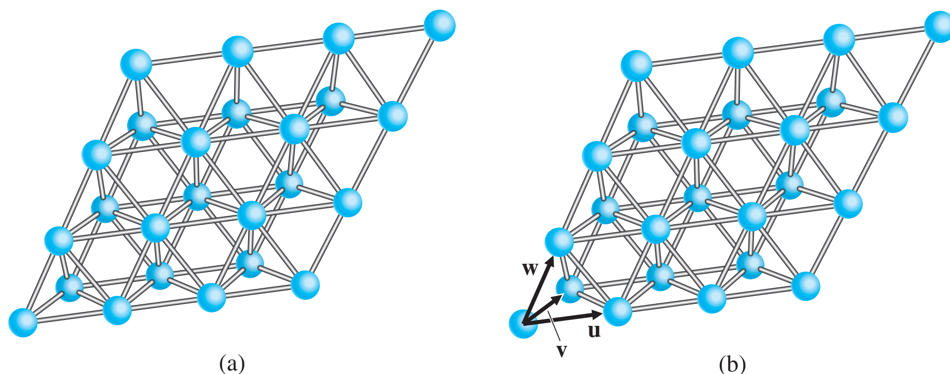


Figura 6.3

Matrices de cambio de base

La figura 6.4 muestra dos sistemas coordenados diferentes para $\mathbb{R}^2$ y cada uno surge de una base diferente. La figura 6.4(a) muestra el sistema coordenado relacionado con la base $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2\}$, mientras que la figura 6.4(b) surge de la base $\mathcal{C} = \{\mathbf{v}_1, \mathbf{v}_2\}$, donde

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

El mismo vector $\mathbf{x}$ se muestra en relación con cada sistema coordenado. Es claro a partir de los diagramas que los vectores coordenados de $\mathbf{x}$ con respecto a $\mathcal{B}$ y $\mathcal{C}$ son

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \quad \text{y} \quad [\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$$

respectivamente. Se evidencia que existe una conexión directa entre los dos vectores coordenados. Una forma de encontrar la relación es usar $[\mathbf{x}]_{\mathcal{B}}$ para calcular

$$\mathbf{x} = \mathbf{u}_1 + 3\mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} + 3\begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$$

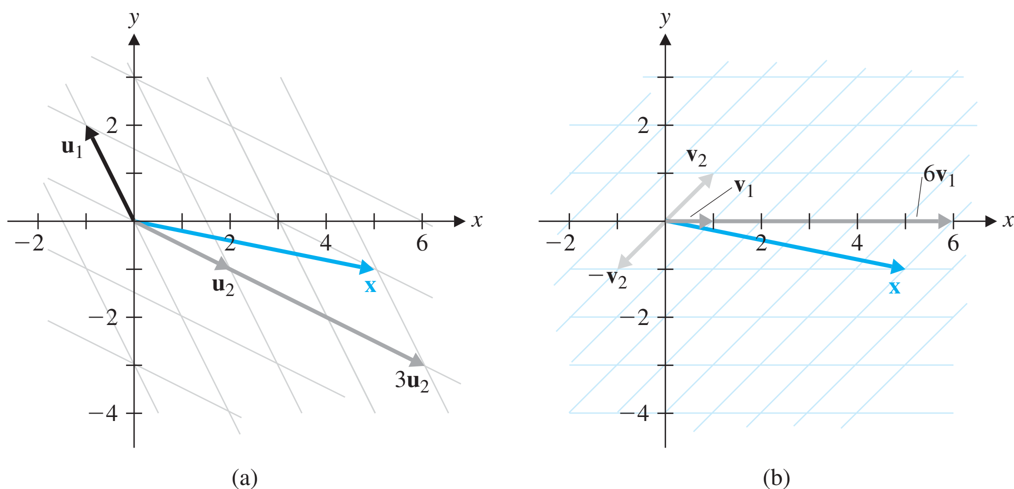


Figura 6.4

Entonces puede encontrar $[\mathbf{x}]_{\mathcal{C}}$ al escribir $\mathbf{x}$ como una combinación lineal de $\mathbf{v}_1$ y $\mathbf{v}_2$. Sin embargo, existe una mejor forma de proceder, una que proporcionará un mecanismo general para tales problemas. Este planteamiento se ilustra en el siguiente ejemplo.

Ejemplo 6.45

Con las anteriores bases $\mathcal{B}$ y $\mathcal{C}$, encuentre $[\mathbf{x}]_{\mathcal{C}}$, dado que $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$.

Solución Puesto que $\mathbf{x} = \mathbf{u}_1 + 3\mathbf{u}_2$, escribir $\mathbf{u}_1$ y $\mathbf{u}_2$ en términos de $\mathbf{v}_1$ y $\mathbf{v}_2$ dará las coordenadas requeridas de $\mathbf{x}$ con respecto a $\mathcal{C}$. Se encuentra que

$$\mathbf{u}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} = -3\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = -3\mathbf{v}_1 + 2\mathbf{v}_2$$

y

$$\mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 3\mathbf{v}_1 - \mathbf{v}_2$$

de modo que

$$\begin{aligned} \mathbf{x} &= \mathbf{u}_1 + 3\mathbf{u}_2 \\ &= (-3\mathbf{v}_1 + 2\mathbf{v}_2) + 3(3\mathbf{v}_1 - \mathbf{v}_2) \\ &= 6\mathbf{v}_1 - \mathbf{v}_2 \end{aligned}$$

Esto produce

$$[\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix} 6 \\ -1 \end{bmatrix}$$

lo que concuerda con la figura 6.4(b).

Este método puede no parecer más sencillo que el sugerido antes del ejemplo 6.45, pero tiene una gran ventaja. Ahora puede encontrar $[\mathbf{y}]_{\mathcal{C}}$ a partir de $[\mathbf{y}]_{\mathcal{B}}$ para *cualquier*

vector $\mathbf{y}$ en $\mathbb{R}^2$ con muy poco trabajo adicional. Observe los cálculos del ejemplo 6.45 desde un punto de vista diferente. A partir de $\mathbf{x} = \mathbf{u}_1 + 3\mathbf{u}_2$, se tiene

$$[\mathbf{x}]_{\mathcal{C}} = [\mathbf{u}_1 + 3\mathbf{u}_2]_{\mathcal{C}} = [\mathbf{u}_1]_{\mathcal{C}} + 3[\mathbf{u}_2]_{\mathcal{C}}$$

por el Teorema 6.6. En consecuencia,

$$\begin{aligned}[\mathbf{x}]_{\mathcal{C}} &= [[\mathbf{u}_1]_{\mathcal{C}}\,[\mathbf{u}_2]_{\mathcal{C}}]\begin{bmatrix}1\\3\end{bmatrix}\\ &= \begin{bmatrix}-3 & 3\\ 2 & -1\end{bmatrix}\begin{bmatrix}1\\3\end{bmatrix}\\ &= P[\mathbf{x}]_{\mathcal{B}}\end{aligned}$$

donde P es la matriz cuyas columnas son $[\mathbf{u}_1]_{\mathcal{C}}$ y $[\mathbf{u}_2]_{\mathcal{C}}$. Este procedimiento se generaliza bastante bien.

Definición Sean $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ y $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ bases para un espacio vectorial V. La matriz de $n \times n$ cuyas columnas son los vectores coordenados $[\mathbf{u}_1]_{\mathcal{C}}, \ldots, [\mathbf{u}_n]_{\mathcal{C}}$ de los vectores en $\mathcal{B}$ con respecto a $\mathcal{C}$ se denota $P_{\mathcal{C}\leftarrow\mathcal{B}}$ y se llama ***matriz de cambio de base*** $\mathcal{B}$ a $\mathcal{C}$. Esto es,

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = [[\mathbf{u}_1]_{\mathcal{C}}\,[\mathbf{u}_2]_{\mathcal{C}}\ \cdots\ [\mathbf{u}_n]_{\mathcal{C}}]$$

Piense en $\mathcal{B}$ como la base "antigua" y en $\mathcal{C}$ como la base "nueva". Entonces las columnas de $P_{\mathcal{C}\leftarrow\mathcal{B}}$ son justo los vectores coordenados obtenidos al escribir los vectores de la base antigua en términos de los nuevos. El Teorema 6.12 muestra que el ejemplo 6.45 es un caso especial de un resultado general.

Teorema 6.12

Sean $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ y $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ bases para un espacio vectorial V y sea $P_{\mathcal{C}\leftarrow\mathcal{B}}$ la matriz de cambio de base $\mathcal{B}$ a $\mathcal{C}$. Entonces

a. $P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$ para todo $\mathbf{x}$ en V.
b. $P_{\mathcal{C}\leftarrow\mathcal{B}}$ es la matriz única P con la propiedad de que $P[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$ para todo $\mathbf{x}$ en V.
c. $P_{\mathcal{C}\leftarrow\mathcal{B}}$ es invertible y $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$.

Demostración (a) Sea $\mathbf{x}$ en V y sea

$$[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}$$

Esto es, $\mathbf{x} = c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n$. Entonces

$$\begin{aligned}[\mathbf{x}]_{\mathcal{C}} &= [c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n]_{\mathcal{C}}\\ &= c_1[\mathbf{u}_1]_{\mathcal{C}} + \cdots + c_n[\mathbf{u}_n]_{\mathcal{C}}\\ &= [[\mathbf{u}_1]_{\mathcal{C}}\ \cdots\ [\mathbf{u}_n]_{\mathcal{C}}]\begin{bmatrix}c_1\\ \vdots\\ c_n\end{bmatrix}\\ &= P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}\end{aligned}$$

(b) Suponga que P es una matriz de $n \times n$ con la propiedad de que $P[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$ para todo $\mathbf{x}$ en V. Al tomar $\mathbf{x} = \mathbf{u}_i$, el i-ésimo vector base en $\mathcal{B}$, se ve que $[\mathbf{x}]_{\mathcal{B}} = [\mathbf{u}_i]_{\mathcal{B}} = \mathbf{e}_i$, de modo que la i-ésima columna de P es

$$\mathbf{p}_i = P\mathbf{e}_i = P[\mathbf{u}_i]_{\mathcal{B}} = [\mathbf{u}_i]_{\mathcal{C}}$$

que es la i-ésima columna de $P_{\mathcal{C}\leftarrow\mathcal{B}}$, por definición. Se tiene que $P = P_{\mathcal{C}\leftarrow\mathcal{B}}$.

(c) Dado que $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ es linealmente independiente en V, el conjunto $\{[\mathbf{u}_1]_{\mathcal{C}}, \ldots, [\mathbf{u}_n]_{\mathcal{C}}\}$ es linealmente independiente en $\mathbb{R}^n$, por el Teorema 6.7. En consecuencia, $P_{\mathcal{C}\leftarrow\mathcal{B}} = [[\mathbf{u}_1]_{\mathcal{C}} \;\cdots\; [\mathbf{u}]_{\mathcal{C}}]$ es invertible, por el teorema fundamental.

Para todo $\mathbf{x}$ en V, se tiene $P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} = [\mathbf{x}]_{\mathcal{C}}$. Al resolver para $[\mathbf{x}]_{\mathcal{B}}$, se encuentra que

$$[\mathbf{x}]_{\mathcal{B}} = (P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}[\mathbf{x}]_{\mathcal{C}}$$

para todo $\mathbf{x}$ en V. Por tanto, $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1}$ es una matriz que cambia bases de $\mathcal{C}$ a $\mathcal{B}$. Por tanto, por la propiedad de unicidad (b), debe tener $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$.

Comentarios

- Puede encontrar útil considerar al cambio de base como una transformación (de hecho, es una transformación lineal) de $\mathbb{R}^n$ a sí mismo que simplemente cambia de un sistema coordenado a otro. La correspondiente transformación a $P_{\mathcal{C}\leftarrow\mathcal{B}}$ acepta $[\mathbf{x}]_{\mathcal{B}}$ como entrada y regresa $[\mathbf{x}]_{\mathcal{C}}$ como salida; $(P_{\mathcal{C}\leftarrow\mathcal{B}})^{-1} = P_{\mathcal{B}\leftarrow\mathcal{C}}$ hace justo lo opuesto. La figura 6.5 ofrece una representación esquemática del proceso.

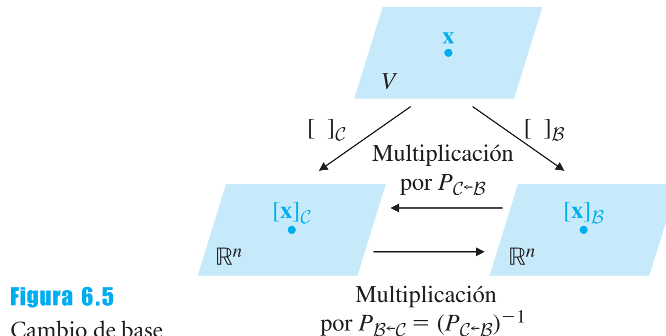


Figura 6.5
Cambio de base

- Las columnas de $P_{\mathcal{C}\leftarrow\mathcal{B}}$ son los vectores coordenados de una base con respecto a la otra base. Para recordar cuál base es cuál, piense en la notación $\mathcal{C} \leftarrow \mathcal{B}$ como decir "$\mathcal{B}$ en términos de $\mathcal{C}$". También es útil recordar que $P_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{x}]_{\mathcal{B}}$ es una combinación lineal de las columnas de $P_{\mathcal{C}\leftarrow\mathcal{B}}$. Pero, dado que el resultado de esta combinación es $[\mathbf{x}]_{\mathcal{C}}$, las columnas de $P_{\mathcal{C}\leftarrow\mathcal{B}}$ deben ser los vectores coordenados con respecto a $\mathcal{C}$.

Ejemplo 6.46

Encuentre las matrices de cambio de base $P_{\mathcal{C}\leftarrow\mathcal{B}}$ y $P_{\mathcal{B}\leftarrow\mathcal{C}}$ para las bases $\mathcal{B} = \{1, x, x^2\}$ y $\mathcal{C} = \{1 + x, x + x^2, 1 + x^2\}$ de $\mathscr{P}_2$. Luego encuentre el vector coordenado de $p(x) = 1 + 2x - x^2$ con respecto a $\mathcal{C}$.

Solución Cambiar una base estándar es sencillo, de modo que primero se encuentra $P_{\mathcal{B}\leftarrow\mathcal{C}}$. Observe que los vectores coordenados para $\mathcal{C}$ en términos de $\mathcal{B}$ son

$$[1 + x]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad [x + x^2]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad [1 + x^2]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

(Vuelva a leer el Comentario después del ejemplo 6.26.) Se tiene que

$$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$



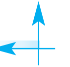
Para encontrar $P_{\mathcal{C}\leftarrow\mathcal{B}}$, podría expresar cada vector en $\mathcal{B}$ como una combinación lineal de los vectores en $\mathcal{C}$ (hágalo), pero es mucho más fácil usar el hecho de que $P_{\mathcal{C}\leftarrow\mathcal{B}} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1}$, por el Teorema 6.12(c). Se encuentra que

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

Ahora se tiene que

$$\begin{aligned} [p(x)]_{\mathcal{C}} &= P_{\mathcal{C}\leftarrow\mathcal{B}}[p(x)]_{\mathcal{B}} \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix} \end{aligned}$$

lo que concuerda con el ejemplo 6.37.



Comentario Si no se necesita $P_{\mathcal{C}\leftarrow\mathcal{B}}$ de manera explícita, puede encontrar $[p(x)]_{\mathcal{C}}$ a partir de $[p(x)]_{\mathcal{B}}$ y $P_{\mathcal{B}\leftarrow\mathcal{C}}$ usando eliminación gaussiana. La reducción por renglón produce

$$[P_{\mathcal{B}\leftarrow\mathcal{C}} \mid [p(x)]_{\mathcal{B}}] \longrightarrow [I \mid (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1}[p(x)]_{\mathcal{B}}] = [I \mid P_{\mathcal{C}\leftarrow\mathcal{B}}[p(x)]_{\mathcal{B}}] = [I \mid [p(x)]_{\mathcal{C}}]$$

(Vea la siguiente sección acerca del uso de la eliminación de Gauss-Jordan.)

Vale la pena repetir el comentario del ejemplo 6.46: cambiar a una base estándar es sencillo. Si $\mathcal{E}$ es la base estándar para un espacio vectorial V y $\mathcal{B}$ es cualquiera otra base, entonces las columnas de $P_{\mathcal{E}\leftarrow\mathcal{B}}$ son los vectores coordenados de $\mathcal{B}$ con respecto a $\mathcal{E}$, y ellas por lo general son "visibles". En el siguiente ejemplo se usa nuevamente este comentario.

Ejemplo 6.47

En M_{22}, sea $\mathcal{B}$ la base $\{E_{11}, E_{21}, E_{12}, E_{22}\}$ y sea $\mathcal{C}$ la base $\{A, B, C, D\}$, donde

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Encuentre la matriz de cambio de base $P_{\mathcal{C}\leftarrow\mathcal{B}}$ y verifique que $[X]_{\mathcal{C}} = P_{\mathcal{C}\leftarrow\mathcal{B}}[X]_{\mathcal{B}}$ para $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$.

Solución 1 Para resolver este problema directamente, debe encontrar los vectores coordenados de $\mathcal{B}$ con respecto a $\mathcal{C}$. Esto involucra resolver cuatro problemas de combinación lineal de la forma $X = aA + bB + cC + dD$, donde X está en $\mathcal{B}$ y debe encontrar a, b, c y d. Sin embargo, aquí se tiene suerte, pues puede encontrar los coeficientes requeridos por inspección.

Claramente, $E_{11} = A$, $E_{21} = -B + C$, $E_{12} = -A + B$ y $E_{22} = -C + D$. Por tanto,

$$[E_{11}]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad [E_{21}]_{\mathcal{C}} = \begin{bmatrix} 0 \\ -1 \\ 1 \\ 0 \end{bmatrix}, \quad [E_{12}]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad [E_{22}]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

de modo que $P_{\mathcal{C}\leftarrow\mathcal{B}} = [[E_{11}]_{\mathcal{C}} \quad [E_{21}]_{\mathcal{C}} \quad [E_{12}]_{\mathcal{C}} \quad [E_{22}]_{\mathcal{C}}] = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Si $X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, entonces

$$[X]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix}$$

y
$$P_{\mathcal{C}\leftarrow\mathcal{B}}[X]_{\mathcal{B}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ 4 \end{bmatrix}$$

Éste es el vector coordenado con respecto a $\mathcal{C}$ de la matriz

$$\begin{aligned} -A - B - C + 4D &= -\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + 4\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = X \end{aligned}$$

como debe ser.

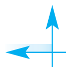
Solución 2 Puede calcular $P_{\mathcal{C}\leftarrow\mathcal{B}}$ en una forma diferente, del modo siguiente. Como se le pedirá probar en el ejercicio 21, si $\mathcal{E}$ es otra base para M_{22}, entonces $P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{C}\leftarrow\mathcal{E}}P_{\mathcal{E}\leftarrow\mathcal{B}} = (P_{\mathcal{E}\leftarrow\mathcal{C}})^{-1}P_{\mathcal{E}\leftarrow\mathcal{B}}$. Si $\mathcal{E}$ es la base estándar, entonces $P_{\mathcal{E}\leftarrow\mathcal{B}}$ y $P_{\mathcal{E}\leftarrow\mathcal{C}}$ pueden encontrarse por inspección. Se tiene

$$P_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad P_{\mathcal{E}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$


(¿Ve por qué?) Por tanto,

$$
\begin{aligned}
P_{\mathcal{C}\leftarrow\mathcal{B}} &= (P_{\mathcal{E}\leftarrow\mathcal{C}})^{-1}P_{\mathcal{E}\leftarrow\mathcal{B}} \\
&= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

lo que concuerda con la primera solución.



Comentario El segundo método tiene la ventaja de no necesitar el cálculo de algunas combinaciones lineales. Tiene la desventaja de requerir que se encuentre un inverso matricial. Sin embargo, usar un CAS facilitará el hallazgo de dicho inverso, de modo que, en general, el segundo método es preferible al primero. No obstante, para ciertos problemas, el primer método puede usarse con igual facilidad. En cualquier caso, está a punto de describirse un tercer método, que puede resultar ser el mejor de todos.

El método de Gauss-Jordan para calcular una matriz de cambio de base

Encontrar la matriz de cambio de basa a una base estándar es sencillo y puede hacerse por inspección. Encontrar la matriz de cambio de base a partir de una base estándar es casi igual de sencillo, pero requiere el cálculo de una matriz inversa, como en el ejemplo 6.46. Si se hace a mano, entonces (excepto para el caso de 2×2) por lo general se encontrará la inversa necesaria mediante eliminación de Gauss-Jordan. Ahora observe una modificación del método de Gauss-Jordan que puede usarse para encontrar la matriz de cambio de base entre dos bases no estándar, como en el ejemplo 6.47.

Suponga que $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ y $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ son bases para un espacio vectorial V y $P_{\mathcal{C}\leftarrow\mathcal{B}}$ es la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$. La i-ésima columna de P es

$$[\mathbf{u}_i]_{\mathcal{C}} = \begin{bmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{bmatrix}$$

de modo que $\mathbf{u}_i = p_{1i}\mathbf{v}_1 + \cdots + p_{ni}\mathbf{v}_n$. Si $\mathcal{E}$ es cualquier base para V, entonces

$$[\mathbf{u}_i]_{\mathcal{E}} = [p_{1i}\mathbf{v}_1 + \cdots + p_{ni}\mathbf{v}_n]_{\mathcal{E}} = p_{1i}[\mathbf{v}_1]_{\mathcal{E}} + \cdots + p_{ni}[\mathbf{v}_n]_{\mathcal{E}}$$

Esto puede reescribirse en forma matricial como

$$\begin{bmatrix} [\mathbf{v}_1]_{\mathcal{E}} & \cdots & [\mathbf{v}_n]_{\mathcal{E}} \end{bmatrix} \begin{bmatrix} p_{1i} \\ \vdots \\ p_{ni} \end{bmatrix} = [\mathbf{u}_i]_{\mathcal{E}}$$

que ahora puede resolverse al aplicar la eliminación de Gauss-Jordan a la matriz aumentada

$$[[\mathbf{v}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{v}_n]_{\mathcal{E}} \mid [\mathbf{u}_i]_{\mathcal{E}}]$$

Existen n de tales sistemas de ecuaciones por resolver, una para cada columna de $P_{\mathcal{C}\leftarrow\mathcal{B}}$, pero *la matriz de coeficientes* $[[\mathbf{v}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{v}_n]_{\mathcal{E}}]$ *es la misma en cada caso*. Por tanto, puede resolver todos los sistemas simultáneamente al reducir por renglón la matriz aumentada de $n \times 2n$

$$[[\mathbf{v}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{v}_n]_{\mathcal{E}} \mid [\mathbf{u}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{u}_n]_{\mathcal{E}}] = [C \mid B]$$

Dado que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es linealmente independiente, también lo es $\{[\mathbf{v}_1]_{\mathcal{E}}, \ldots, [\mathbf{v}_n]_{\mathcal{E}}\}$ por el Teorema 6.7. Por tanto, la matriz C cuyas columnas son $\{[\mathbf{v}_1]_{\mathcal{E}}, \ldots, [\mathbf{v}_n]_{\mathcal{E}}\}$ tiene la matriz identidad I de $n \times n$ para su forma escalonada reducida por renglón, por el teorema fundamental. Se tiene que la eliminación de Gauss-Jordan necesariamente producirá

$$[C \mid B] \rightarrow [I \mid P]$$

donde $P = P_{\mathcal{C}\leftarrow\mathcal{B}}$.

Se acaba de demostrar el siguiente teorema.

Teorema 6.13 Sean $\mathcal{B} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ y $\mathcal{C} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ bases para un espacio vectorial V. Sean $B = [[\mathbf{u}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{u}_n]_{\mathcal{E}}]$ y $C = [[\mathbf{v}_1]_{\mathcal{E}} \ \cdots \ [\mathbf{v}_n]_{\mathcal{E}}]$, donde $\mathcal{E}$ es cualquier base para V. Entonces la reducción por renglón aplicada a la matriz aumentada $[C \mid B]$ de $n \times 2n$ produce

$$[C \mid B] \rightarrow [I \mid P_{\mathcal{C}\leftarrow\mathcal{B}}]$$

Si $\mathcal{E}$ es una base estándar, este método es particularmente fácil de usar, pues en dicho caso $B = P_{\mathcal{E}\leftarrow\mathcal{B}}$ y $C = P_{\mathcal{E}\leftarrow\mathcal{C}}$. Este método se ilustra al volver a trabajar el problema del ejemplo 6.47.

Ejemplo 6.48 Vuelva a resolver el ejemplo 6.47 usando el método de Gauss-Jordan.

Solución Al tomar $\mathcal{E}$ como la base estándar para M_{22}, se ve que

$$B = P_{\mathcal{E}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad C = P_{\mathcal{E}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

La reducción por renglón produce

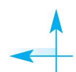
$$[C \mid B] = \left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right]$$



(Verifique esta reducción por renglón.) Se tiene que

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

como se encontró anteriormente.



Ejercicios 6.3

En los ejercicios 1-4:
(a) Encuentre los vectores coordenados $[\mathbf{x}]_{\mathcal{B}}$ *y* $[\mathbf{x}]_{\mathcal{C}}$ *de* $\mathbf{x}$ *con respecto a las bases* $\mathcal{B}$ *y* $\mathcal{C}$, *respectivamente.*
(b) Encuentre la matriz de cambio de base $P_{\mathcal{C}\leftarrow\mathcal{B}}$ *de* $\mathcal{B}$ *a* $\mathcal{C}$.
(c) Use su respuesta al inciso (b) para calcular $[\mathbf{x}]_{\mathcal{C}}$, *y compare su respuesta con la que encontró en el inciso (a).*
(d) Encuentre la matriz de cambio de base $P_{\mathcal{B}\leftarrow\mathcal{C}}$ *de* $\mathcal{C}$ *a* $\mathcal{B}$.
(e) Use sus respuestas a los incisos (c) y (d) para calcular $[\mathbf{x}]_{\mathcal{B}}$, *y compare su respuesta con la que encontró en el inciso (a).*

1. $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ en $\mathbb{R}^2$

2. $\mathbf{x} = \begin{bmatrix} 3 \\ -4 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ en $\mathbb{R}^2$

3. $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ en $\mathbb{R}^3$

4. $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \mathcal{B} = \left\{ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ en $\mathbb{R}^3$

En los ejercicios 5-8, siga las instrucciones para los ejercicios 1-4 y use $p(x)$ *en lugar de* $\mathbf{x}$.

5. $p(x) = 2 - x, \mathcal{B} = \{1, x\}, \mathcal{C} = \{x, 1 + x\}$ en $\mathscr{P}_1$

6. $p(x) = 3 + 2x, \mathcal{B} = \{1 + x, 1 - x\},$
$\mathcal{C} = \{2x, 3\}$ en $\mathscr{P}_1$

7. $p(x) = 1 + x^2, \mathcal{B} = \{1 + x + x^2, x + x^2, x^2\},$
$\mathcal{C} = \{1, x, x^2\}$ en $\mathscr{P}_2$

8. $p(x) = 2 - x + 3x^2, \mathcal{B} = \{1 + x, 1 + x^2, x + x^2\},$
$\mathcal{C} = \{1, x, 1 + x + x^2\}$ en $\mathscr{P}_2$

En los ejercicios 9 y 10, siga las instrucciones para los ejercicios 1-4 y use A *en lugar de* $\mathbf{x}$.

9. $A = \begin{bmatrix} 4 & 2 \\ 0 & -1 \end{bmatrix}$, $\mathcal{B}$ = la base estándar,
$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\}$ en M_{22}

10. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$
$\mathcal{B} = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \right\},$
$\mathcal{C} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$ en M_{22}

En los ejercicios 11 y 12, siga las instrucciones para los ejercicios 1-4 y use $f(x)$ *en lugar de* $\mathbf{x}$.

11. $f(x) = 2 \operatorname{sen} x - 3 \cos x, \mathcal{B} = \{\operatorname{sen} x + \cos x, \cos x\},$
$\mathcal{C} = \{\operatorname{sen} x + \cos x, \operatorname{sen} x - \cos x\}$ en gen$(\operatorname{sen} x, \cos x)$

12. $f(x) = \cos x, \mathcal{B} = \{\operatorname{sen} x, \operatorname{sen} x + \cos x\},$
$\mathcal{C} = \{\operatorname{sen} x + \cos x, \operatorname{sen} x - \cos x\}$ en gen$(\operatorname{sen} x, \cos x)$

13. Rote el eje xy en el plano en sentido contrario al de las manecillas del reloj un ángulo $\theta = 60°$ para obtener nuevos ejes $x'y'$. Use los métodos de esta sección para encontrar (a) las coordenadas $x'y'$ del punto cuyas coordenadas xy son $(3, 2)$ y (b) las coordenadas xy del punto cuyas coordenadas $x'y'$ son $(4, -4)$.

14. Repita el ejercicio 13 con $\theta = 135°$.

15. Sean $\mathcal{B}$ y $\mathcal{C}$ bases para $\mathbb{R}^2$. Si $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$ y la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

encuentre $\mathcal{B}$.

16. Sean $\mathcal{B}$ y $\mathcal{C}$ bases para $\mathscr{P}_2$. Si $\mathcal{B} = \{x, 1 + x, 1 - x + x^2\}$ y la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es

$$P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

encuentre $\mathcal{C}$.

En cálculo se aprende que un ***polinomio de Taylor de grado n en torno a a*** *es un polinomio de la forma*

$p(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \cdots + a_n(x - a)^n$ *donde* $a_n \neq 0$. *En otras palabras, es un polinomio que se expandió en términos de potencias de x – a en lugar de potencias de x. Los polinomios de Taylor son muy útiles para aproximar funciones que "se comportan bien" cerca de x = a.*

El conjunto $\mathcal{B} = \{1, x - a, (x - a)^2, \ldots, (x - a)^n\}$ *es una base para* $\mathcal{P}_n$ *para cualquier número real a. (¿Percibe una forma rápida de demostrar esto? Intente usar el Teorema 6.7.) Este hecho permite usar las técnicas de esta sección para reescribir un polinomio como un polinomio de Taylor en torno a una a dada.*

17. Exprese $p(x) = 1 + 2x - 5x^2$ como un polinomio de Taylor en torno a $a = 1$.

18. Exprese $p(x) = 1 + 2x - 5x^2$ como un polinomio de Taylor en torno a $a = -2$.

19. Exprese $p(x) = x^3$ como un polinomio de Taylor en torno a $a = -1$.

20. Exprese $p(x) = x^3$ como un polinomio de Taylor en torno a $a = \frac{1}{2}$.

21. Sean $\mathcal{B}, \mathcal{C}$ y $\mathcal{D}$ bases para un espacio vectorial V de dimensión finita. Pruebe que

$$P_{\mathcal{D}\leftarrow\mathcal{C}}P_{\mathcal{C}\leftarrow\mathcal{B}} = P_{\mathcal{D}\leftarrow\mathcal{B}}$$

22. Sea V un espacio vectorial de dimensión n con base $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Sea una matriz invertible de $n \times n$ y establezca

$$\mathbf{u}_i = p_{1i}\mathbf{v}_1 + \cdots + p_{ni}\mathbf{v}_n$$

para $i = 1, \ldots, n$. Pruebe que $\mathcal{C} = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ es una base para V y demuestre que $P = P_{\mathcal{B}\leftarrow\mathcal{C}}$.

6.4 Transformaciones lineales

En la sección 3.6 se encontraron las transformaciones lineales en el contexto de las transformaciones matriciales de $\mathbb{R}^n$ a $\mathbb{R}^m$. En esta sección se amplía este concepto a transformaciones lineales entre espacios vectoriales arbitrarios.

Definición Una ***transformación lineal*** de un espacio vectorial V a un espacio vectorial W es un mapeo $T: V \to W$ tal que, para todo $\mathbf{u}$ y $\mathbf{v}$ en V y para todos los escalares c,

1. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$
2. $T(c\mathbf{u}) = cT(\mathbf{u})$

Es directo demostrar que esta definición es equivalente al requisito de que T conserve todas las combinaciones lineales. Esto es,

$T: V \to W$ es una transformación lineal si y sólo si

$$T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_kT(\mathbf{v}_k)$$

para todo $\mathbf{v}_1, \ldots, \mathbf{v}_k$ en V y escalares $c_1, \ldots, c_k$.

Ejemplo 6.49 Toda transformación matricial es una transformación lineal. Esto es, si A es una matriz de $m \times n$, entonces la transformación $T_A: \mathbb{R}^n \to \mathbb{R}^m$ definida por

$$T_A(\mathbf{x}) = A\mathbf{x} \text{ para } \mathbf{x} \text{ en } \mathbb{R}^n$$

es una transformación lineal. Éste es un nuevo planteamiento del Teorema 3.30.

Ejemplo 6.50

Defina $T: M_{nn} \to M_{nn}$ mediante $T(A) = A^T$. Demuestre que T es una transformación lineal.

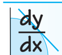
Solución Se comprueba que, para A y B en M_{nn} y escalares c,

$$T(A + B) = (A + B)^T = A^T + B^T = T(A) + T(B)$$

y

$$T(cA) = (cA)^T = cA^T = cT(A)$$

Por tanto, T es una transformación lineal.



Ejemplo 6.51

Sea D el ***operador diferencial*** $D: \mathscr{D} \to \mathscr{F}$ definido por $D(f) = f'$. Demuestre que D es una transformación lineal.

Solución Sean f y g funciones derivables y sea c un escalar. Entonces, a partir del cálculo, se sabe que

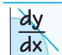
$$D(f + g) = (f + g)' = f' + g' = D(f) + D(g)$$

y

$$D(cf) = (cf)' = cf' = cD(f)$$

Por tanto, D es una transformación lineal.

En cálculo se aprende que toda función continua en $[a, b]$ es integrable. El siguiente ejemplo demuestra que la integración es una transformación lineal.



Ejemplo 6.52

Defina S: $\mathscr{C}[a, b] \to \mathbb{R}$ mediante $S(f) = \int_a^b f(x)dx$. Demuestre que S es una transformación lineal.

Solución Sean f y g en $\mathscr{C}[a, b]$. Entonces

$$\begin{aligned} S(f + g) &= \int_a^b (f + g)(x)\, dx \\ &= \int_a^b (f(x) + g(x))\, dx \\ &= \int_a^b f(x)\, dx + \int_a^b g(x)\, dx \\ &= S(f) + S(g) \end{aligned}$$

y

$$\begin{aligned} S(cf) &= \int_a^b (cf)(x)\, dx \\ &= \int_a^b cf(x)\, dx \\ &= c\int_a^b f(x)\, dx \\ &= cS(f) \end{aligned}$$

Se tiene que S es lineal.

Ejemplo 6.53

Demuestre que ninguna de las siguientes transformaciones es lineal:

(a) $T : M_{22} \to \mathbb{R}$ definida por $T(A) = \det A$

(b) $T : \mathbb{R} \to \mathbb{R}$ definida por $T(x) = 2^x$

(c) $T : \mathbb{R} \to \mathbb{R}$ definida por $T(x) = x + 1$

Solución En cada caso se ofrece un contraejemplo específico para demostrar que una de las propiedades de una transformación lineal no se cumple.

(a) Sean $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ y $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Entonces $A + B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, de modo que

$$T(A + B) = \det(A + B) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

Pero

$$T(A) + T(B) = \det A + \det B = \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} 0 & 0 \\ 0 & 1 \end{vmatrix} = 0 + 0 = 0$$

de modo que $T(A + B) \neq T(A) + T(B)$ y T no es lineal.

(b) Sean $x = 1$ y $y = 2$. Entonces

$$T(x + y) = T(3) = 2^3 = 8 \neq 6 = 2^1 + 2^2 = T(x) + T(y)$$

de modo que T no es lineal.

(c) Sean $x = 1$ y $y = 2$. Entonces

$$T(x + y) = T(3) = 3 + 1 = 4 \neq 5 = (1 + 1) + (2 + 1) = T(x) + T(y)$$

Por tanto, T no es lineal.

Comentario El ejemplo 6.53(c) demuestra que es necesario tener cuidado cuando uno encuentra la palabra "lineal". Como *función*, $T(x) = x + 1$ es lineal, pues su gráfica es una recta. Sin embargo, no es una *transformación lineal* del espacio vectorial $\mathbb{R}$ por sí misma, pues falla en satisfacer la definición. (¿Qué funciones lineales de $\mathbb{R}$ a $\mathbb{R}$ también serán transformaciones lineales?)

Existen dos transformaciones lineales especiales que merecen distinguirse.

Ejemplo 6.54

(a) Para cualesquier espacios vectoriales V y W, la transformación $T_0 : V \to W$ que mapea todo vector en V al vector cero en W se llama ***transformación cero***. Esto es,

$$T_0(\mathbf{v}) = \mathbf{0} \quad \text{para todo } \mathbf{v} \text{ en } V$$

(b) Para cualquier espacio vectorial V, la transformación $I : V \to V$ que mapea todo vector en V a sí mismo se llama ***transformación identidad***. Esto es,

$$I(\mathbf{v}) = \mathbf{v} \quad \text{para todo } \mathbf{v} \text{ en } V$$

(Si es importante identificar el espacio vectorial V, puede escribir I_V por claridad.) Las demostraciones de que las transformaciones cero e identidad son lineales se dejan como ejercicios sencillos.

Propiedades de las transformaciones lineales

En el capítulo 3, todas las transformaciones lineales eran transformaciones matriciales, y sus propiedades se relacionaban directamente con propiedades de las matrices involucradas. El siguiente teorema es fácil de demostrar para transformaciones matriciales. (¡Hágalo!) La demostración completa para transformaciones lineales en general requiere un poco más de cuidado, pero todavía es directa.

Teorema 6.14 Sea $T: V \to W$ una transformación lineal. Entonces:

a. $T(\mathbf{0}) = \mathbf{0}$
b. $T(-\mathbf{v}) = -T(\mathbf{v})$ para todo $\mathbf{v}$ en V.
c. $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$ para todo $\mathbf{u}$ y $\mathbf{v}$ en V.

Demostración Se demuestran las propiedades (a) y (c) y se deja la demostración de la propiedad (b) como ejercicio 21.

(a) Sea $\mathbf{v}$ cualquier vector en V. Entonces $T(\mathbf{0}) = T(0\mathbf{v}) = 0T(\mathbf{v}) = \mathbf{0}$, como se requiere. (¿Puede dar una razón para cada paso?)

(c) $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u} + (-1)\mathbf{v}) = T(\mathbf{u}) + (-1)T(\mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v})$

Comentario La propiedad (a) puede ser útil para demostrar que ciertas transformaciones *no* son lineales. Como ilustración, considere el ejemplo 6.53(b). Si $T(x) = 2^x$, entonces $T(0) = 2^0 = 1 \neq 0$, de modo que T no es lineal, por el Teorema 6.14(a). Sin embargo, tenga cuidado, pues existen muchas transformaciones que *sí* mapean el vector cero al vector cero, mas todavía *no* son lineales. El ejemplo 6.53(a) es un caso a la medida: el vector cero es la matriz cero O de 2×2, de modo que $T(O) = \det O = 0$, pero se vio que $T(A) = \det A$ no es lineal.

La propiedad más importante de una transformación lineal $T: V \to W$ es que T está completamente determinada por su efecto sobre una base para V. El siguiente ejemplo demuestra lo que esto significa.

Ejemplo 6.55 Suponga que T es una transformación lineal de $\mathbb{R}^2$ a $\mathscr{P}_2$ tal que

$$T\begin{bmatrix}1\\1\end{bmatrix} = 2 - 3x + x^2 \quad \text{y} \quad T\begin{bmatrix}2\\3\end{bmatrix} = 1 - x^2$$

Encuentre $T\begin{bmatrix}-1\\2\end{bmatrix}$ y $T\begin{bmatrix}a\\b\end{bmatrix}$.

Solución Dado que $\mathcal{B} = \left\{\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}2\\3\end{bmatrix}\right\}$ es una base para $\mathbb{R}^2$ (¿por qué?), todo vector en $\mathbb{R}^2$ está en gen($\mathcal{B}$). Al resolver

$$c_1\begin{bmatrix}1\\1\end{bmatrix} + c_2\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}-1\\2\end{bmatrix}$$

se encuentra que $c_1 = -7$ y $c_2 = 3$. Por tanto,

$$\begin{aligned} T\begin{bmatrix} -1 \\ 2 \end{bmatrix} &= T\left(-7\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) \\ &= -7T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3T\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= -7(2 - 3x + x^2) + 3(1 - x^2) \\ &= -11 + 21x - 10x^2 \end{aligned}$$

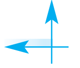
De igual modo, se descubre que

$$\begin{bmatrix} a \\ b \end{bmatrix} = (3a - 2b)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

de modo que

$$\begin{aligned} T\begin{bmatrix} a \\ b \end{bmatrix} &= T\left((3a - 2b)\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) \\ &= (3a - 2b)T\begin{bmatrix} 1 \\ 1 \end{bmatrix} + (b - a)T\begin{bmatrix} 2 \\ 3 \end{bmatrix} \\ &= (3a - 2b)(2 - 3x + x^2) + (b - a)(1 - x^2) \\ &= (5a - 3b) + (-9a + 6b)x + (4a - 3b)x^2 \end{aligned}$$

(Note que, al hacer $a = -1$ y $b = 2$, se recupera la solución $T\begin{bmatrix} -1 \\ 2 \end{bmatrix} = -11 + 21x - 10x^2$.)



La demostración del teorema general es bastante directa.

Teorema 6.14 Sea $T: V \to W$ una transformación lineal y sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto generador para V. Entonces $T(\mathcal{B}) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ genera el rango de T.

Demostración El rango de T es el conjunto de todos los vectores en W que son de la forma $T(\mathbf{v})$, donde $\mathbf{v}$ está en V. Sea $T(\mathbf{v})$ que está en el rango de T. Dado que $\mathcal{B}$ genera a V, existen escalares $c_1, \ldots, c_n$ tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

Al aplicar T y usar el hecho de que es una transformación lineal, se ve que

$$T(\mathbf{v}) = T(c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + \cdots + c_nT(\mathbf{v}_n)$$

En otras palabras, $T(\mathbf{v})$ está en gen$(T(\mathcal{B}))$, como se requiere.

El Teorema 6.15 se aplica, en particular, cuando $\mathcal{B}$ es una base para V. Puede suponer que, en este caso, $T(\mathcal{B})$ sería entonces una base para el rango de T. Por desgracia, este no siempre es el caso. Este conflicto se abordará en la sección 6.5.

Composición de transformaciones lineales

En la sección 3.6 se definió la composición de transformaciones matriciales. La definición se extiende a transformaciones lineales generales en una forma obvia.

$S \circ T$ se lee "S de T."

Definición Si $T: U \to V$ y $S: V \to W$ son transformaciones lineales, entonces la ***composición de S con T*** es el mapeo $S \circ T$, definido por

$$(S \circ T)(\mathbf{u}) = S(T(\mathbf{u}))$$

donde **u** está en U.

Observe que $S \circ T$ es un mapeo de U a W (vea la figura 6.6). Note también que, para que la definición tenga sentido, el rango de T debe estar contenido en el dominio de S.

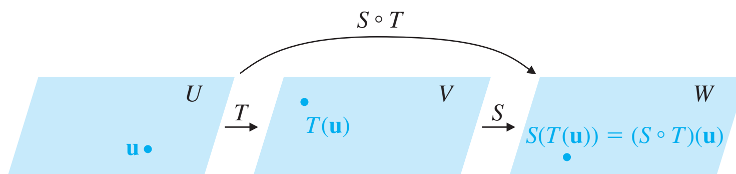


Figura 6.6
Composición de transformaciones lineales

Ejemplo 6.56

Sean $T: \mathbb{R}^2 \to \mathscr{P}_1$ y $S: \mathscr{P}_1 \to \mathscr{P}_2$ transformaciones lineales definidas por

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a+b)x \qquad \text{y} \qquad S(p(x)) = xp(x)$$

Encuentre $(S \circ T)\begin{bmatrix} 3 \\ -2 \end{bmatrix}$ y $(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix}$.

Solución Calcule

$$(S \circ T)\begin{bmatrix} 3 \\ -2 \end{bmatrix} = S\left(T\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right) = S(3 + (3-2)x) = S(3+x) = x(3+x)$$
$$= 3x + x^2$$

y

$$(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix} = S\left(T\begin{bmatrix} a \\ b \end{bmatrix}\right) = S(a + (a+b)x) = x(a + (a+b)x)$$
$$= ax + (a+b)x^2$$

El capítulo 3 demostró que la composición de dos transformaciones matriciales era otra transformación matricial. En general, se tiene el siguiente teorema.

Teorema 6.16 Si $T: U \to V$ y $S: V \to W$ son transformaciones lineales, entonces $S \circ T: U \to W$ es una transformación lineal.

Demostración Sean **u** y **v** que están en U y sea c un escalar. Entonces

$$\begin{aligned}(S \circ T)(\mathbf{u} + \mathbf{v}) &= S(T(\mathbf{u} + \mathbf{v})) \\ &= S(T(\mathbf{u}) + T(\mathbf{v})) && \text{pues } T \text{ es lineal} \\ &= S(T(\mathbf{u})) + S(T(\mathbf{v})) && \text{pues } S \text{ es lineal} \\ &= (S \circ T)(\mathbf{u}) + (S \circ T)(\mathbf{v})\end{aligned}$$

y

$$\begin{aligned}(S \circ T)(c\mathbf{u}) &= S(T(c\mathbf{u})) \\ &= S(cT(\mathbf{u})) && \text{pues } T \text{ es lineal} \\ &= cS(T(\mathbf{u})) && \text{pues } S \text{ es lineal} \\ &= c(S \circ T)(\mathbf{u})\end{aligned}$$

Por tanto, $S \circ T$ es una transformación lineal.

Las propiedades algebraicas de las transformaciones lineales imitan las de las transformaciones matriciales, que, a su vez, se relacionan con las propiedades algebraicas de las matrices. Por ejemplo, la composición de las transformaciones lineales es asociativa. Esto es: si R, S y T son transformaciones lineales, entonces

$$R \circ (S \circ T) = (R \circ S) \circ T$$

siempre que dichas composiciones tengan sentido. La demostración de esta propiedad es idéntica a la dada en la sección 3.6.

El siguiente ejemplo ofrece otra útil propiedad (más no sorprendente) de las transformaciones lineales.

Ejemplo 6.57

Sean $S : U \to V$ y $T : V \to W$ transformaciones lineales y sea $I : V \to V$ la transformación identidad. Entonces, para todo **v** en V, se tiene

$$(T \circ I)(\mathbf{v}) = T(I(\mathbf{v})) = T(\mathbf{v})$$

Dado que $T \circ I$ y T tienen el mismo valor en todo **v** en su dominio, se tiene que $T \circ I = T$. De igual modo, $I \circ S = S$.

Comentario El método del ejemplo 6.57 no vale nada. Suponga que quiere demostrar que dos transformaciones lineales T_1 y T_2 (ambas de V a W) son iguales. Es suficiente con demostrar que $T_1(\mathbf{v}) = T_2(\mathbf{v})$ para todo **v** en V.

En los ejercicios se exploran más propiedades de las transformaciones lineales.

Inversas de transformaciones lineales

Definición Una transformación lineal $T : V \to W$ es ***invertible*** si existe una transformación lineal $T' : W \to V$ tal que

$$T' \circ T = I_V \quad \text{y} \quad T \circ T' = I_W$$

En este caso, T' se llama ***inversa*** de T.

Observaciones

- El dominio V y el codominio W de T no tienen que ser lo mismo, como lo son en el caso de las transformaciones matriciales invertibles. Sin embargo, en la siguiente sección se verá que V y W deben estar estrechamente relacionados.
- El requisito de que T' sea lineal podría omitirse de esta definición. Porque, como se verá en el Teorema 6.24, si T' es *cualquier* mapeo de W a V tal que $T' \circ T = I_V$ y $T \circ T' = I_W$, entonces T' también está forzado a ser lineal.
- Si T' es una inversa de T, entonces la definición implica que T es una inversa de T'. Por tanto, T' también es invertible.

Ejemplo 6.58

Verifique que los mapeos $T: \mathbb{R}^2 \to \mathscr{P}_1$ y $T': \mathscr{P}_1 \to \mathbb{R}^2$ definidos por

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a+b)x \quad \text{y} \quad T'(c+dx) = \begin{bmatrix} c \\ d-c \end{bmatrix}$$

son inversas.

Solución Calcule

$$(T' \circ T)\begin{bmatrix} a \\ b \end{bmatrix} = T'\left(T\begin{bmatrix} a \\ b \end{bmatrix}\right) = T'(a + (a+b)x) = \begin{bmatrix} a \\ (a+b)-a \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

y

$$(T \circ T')(c+dx) = T(T'(c+dx)) = T\begin{bmatrix} c \\ d-c \end{bmatrix} = c + (c + (d-c))x = c + dx$$

Por tanto, $T' \circ T = I_{\mathbb{R}^2}$ y $T \circ T' = I_{\mathscr{P}_1}$. En consecuencia, T y T' son inversas una de la otra.

Como fue el caso con las matrices invertibles, las inversas de las transformaciones lineales son únicas si existen. El siguiente teorema es el análogo del Teorema 3.6.

Teorema 6.17

Si T es una transformación lineal invertible, entonces su inversa es única.

Demostración La demostración es la misma que la del Teorema 3.6, con productos de matrices reemplazadas por composiciones de transformaciones lineales.

Gracias al Teorema 6.17, si T es invertible, puede referírsele como *la* inversa de T. Se le denotará T^{-1} (pronúnciese "T inversa"). En las dos secciones siguientes se abordará el conflicto de determinar cuándo una transformación lineal dada es invertible y encontrar su inversa cuando exista.

Ejercicios 6.4

En los ejercicios 1-12, determine si T es una transformación lineal.

1. $T: M_{22} \to M_{22}$ definida por

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ 0 & c+d \end{bmatrix}$$

2. $T: M_{22} \to M_{22}$ definida por

$$T\begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} w+x & 1 \\ 0 & y-z \end{bmatrix}$$

3. $T: M_{nn} \to M_{nn}$ definida por $T(A) = AB$, donde B es una matriz fija de $n \times n$.

4. $T: M_{nn} \to M_{nn}$ definida por $T(A) = AB - BA$, donde B es una matriz fija de $n \times n$.

5. $T: M_{nn} \to \mathbb{R}$ definida por $T(A) = \text{tr}(A)$

6. $T: M_{nn} \to \mathbb{R}$ definida por $T(A) = a_{11}a_{22} \cdots a_{nn}$

7. $T: M_{nn} \to \mathbb{R}$ definida por $T(A) = \text{rank}(A)$

8. $T: \mathscr{P}_2 \to \mathscr{P}_2$ definida por $T(a + bx + cx^2) = (a+1) + (b+1)x + (c+1)x^2$

9. $T: \mathscr{P}_2 \to \mathscr{P}_2$ definida por $T(a + bx + cx^2) = a + b(x+1) + b(x+1)^2$

10. $T: \mathscr{F} \to \mathscr{F}$ definida por $T(f) = f(x^2)$

11. $T: \mathscr{F} \to \mathscr{F}$ definida por $T(f) = (f(x))^2$

12. $T: \mathscr{F} \to \mathbb{R}$ definida por $T(f) = f(c)$, donde c es un escalar fijo

13. Demuestre que las transformaciones S y T del ejemplo 6.56 son lineales.

14. Sea $T: \mathbb{R}^2 \to \mathbb{R}^3$ una transformación lineal para la cual

$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \text{y} \quad T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

Encuentre $T\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ y $T\begin{bmatrix} a \\ b \end{bmatrix}$.

15. Sea $T: \mathbb{R}^2 \to \mathscr{P}_2$ una transformación lineal para la cual

$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 1 - 2x \quad \text{y} \quad T\begin{bmatrix} 3 \\ -1 \end{bmatrix} = x + 2x^2$$

Encuentre $T\begin{bmatrix} -7 \\ 9 \end{bmatrix}$ y $T\begin{bmatrix} a \\ b \end{bmatrix}$.

16. Sea $T: \mathscr{P}_2 \to \mathscr{P}_1$ una transformación lineal para la cual

$$T(1) = 1 + x, \quad T(x) = 3 - 2x \quad \text{y} \quad T(x^2) = 4 + 3x$$

Encuentre $T(4 - x + 3x^2)$ y $T(a + bx + cx^2)$.

17. Sea $T: \mathscr{P}_2 \to \mathscr{P}_2$ una transformación lineal para la cual

$$T(1 + x) = 1 + x^2, \quad T(x + x^2) = x - x^2,$$
$$T(1 + x^2) = 1 + x + x^2$$

Encuentre $T(4 - x + 3x^2)$ y $T(a + bx + cx^2)$.

18. Sea $T: M_{22} \to \mathbb{R}$ una transformación lineal para la cual

$$T\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = 1, \quad T\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 2,$$
$$T\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = 3, \quad T\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = 4$$

Encuentre $T\begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$ y $T\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

19. Sea $T: M_{22} \to \mathbb{R}$ una transformación lineal. Demuestre que existen escalares a, b, c y d tales que

$$T\begin{bmatrix} w & x \\ y & z \end{bmatrix} = aw + bx + cy + dz$$

para todo $\begin{bmatrix} w & x \\ y & z \end{bmatrix}$ en M_{22}.

20. Demuestre que no existe transformación lineal $T: \mathbb{R}^3 \to \mathscr{P}_2$ tal que

$$T\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = 1 + x, \quad T\begin{bmatrix} 3 \\ 0 \\ 2 \end{bmatrix} = 2 - x + x^2,$$
$$T\begin{bmatrix} 0 \\ 6 \\ -8 \end{bmatrix} = -2 + 2x^2$$

21. Demuestre el Teorema 6.14(b).

22. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base para un espacio vectorial V y sea $T: V \to V$ una transformación lineal. Pruebe que si $T(\mathbf{v}_1) = \mathbf{v}_1, T(\mathbf{v}_2) = \mathbf{v}_2 \ldots, T(\mathbf{v}_n) = \mathbf{v}_n$, entonces T es la transformación identidad en V.


23. Sea $T: \mathscr{P}_n \to \mathscr{P}_n$ una transformación lineal tal que $T(x^k) = kx^{k-1}$ para $k = 0, 1, \ldots, n$. Demuestre que T debe ser el operador diferencial D.

24. Sean $\mathbf{v}_1, \ldots, \mathbf{v}_n$ vectores en un espacio vectorial V y sea $T: V \to W$ una transformación lineal.

(**a**) Si $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ es linealmente independiente en W, demuestre que $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es linealmente independiente en V.

(**b**) Demuestre que lo contrario del inciso (a) es falso. Esto es, no necesariamente es cierto que si $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es linealmente independiente en V, entonces $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ es linealmente independiente en W. Ilustre esto con un ejemplo $T: \mathbb{R}^2 \to \mathbb{R}^2$.

25. Defina las transformaciones lineales $S: \mathbb{R}^2 \to M_{22}$ y $T: \mathbb{R}^2 \to \mathbb{R}^2$ mediante

$$S\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+b & b \\ 0 & a-b \end{bmatrix} \quad \text{y} \quad T\begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} 2c+d \\ -d \end{bmatrix}$$

Calcule $(S \circ T)\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ y $(S \circ T)\begin{bmatrix} x \\ y \end{bmatrix}$. ¿Puede calcular $(T \circ S)\begin{bmatrix} x \\ y \end{bmatrix}$? Si es así, hágalo.

26. Defina las transformaciones lineales $S: \mathscr{P}_1 \to \mathscr{P}_2$ y $T: \mathscr{P}_2 \to \mathscr{P}_1$ mediante

$$S(a + bx) = a + (a+b)x + 2bx^2$$

y $$T(a + bx + cx^2) = b + 2cx$$

Calcule $(S \circ T)(3 + 2x - x^2)$ y $(S \circ T)(a + bx + cx^2)$. ¿Puede calcular $(T \circ S)(a + bx)$? Si es así, hágalo.

27. Defina las transformaciones lineales $S: \mathscr{P}_n \to \mathscr{P}_n$ y $T: \mathscr{P}_n \to \mathscr{P}_n$ mediante

$$S(p(x)) = p(x+1) \quad \text{y} \quad T(p(x)) = p'(x)$$

Encuentre $(S \circ T)(p(x))$ y $(T \circ S)(p(x))$. [*Sugerencia:* recuerde la regla de la cadena.]

28. Defina las transformaciones lineales $S: \mathscr{P}_n \to \mathscr{P}_n$ y $T: \mathscr{P}_n \to \mathscr{P}_n$ mediante

$$S(p(x)) = p(x+1) \quad \text{y} \quad T(p(x)) = xp'(x)$$

Encuentre $(S \circ T)(p(x))$ y $(T \circ S)(p(x))$.

En los ejercicios 29 y 30, verifique que S y T son inversas.

29. $S: \mathbb{R}^2 \to \mathbb{R}^2$ definida por $S\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4x+y \\ 3x+y \end{bmatrix}$ y $T: \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ -3x+4y \end{bmatrix}$

30. $S: \mathscr{P}_1 \to \mathscr{P}_1$ definida por $S(a + bx) = (-4a + b) + 2ax$ y $T: \mathscr{P}_1 \to \mathscr{P}_1$ definida por

$$T(a + bx) = b/2 + (a + 2b)x$$

31. Demuestre el Teorema 6.17.

32. Sea $T: V \to V$ una transformación lineal tal que $T \circ T = I$.

(**a**) Demuestre que $\{\mathbf{v}, T(\mathbf{v})\}$ es linealmente dependiente y sólo si $T(\mathbf{v}) = \pm\mathbf{v}$.

(**b**) Ofrezca un ejemplo de tal transformación lineal con $V = \mathbb{R}^2$.

33. Sea $T: V \to V$ una transformación lineal tal que $T \circ T = T$.

(**a**) Demuestre que $\{\mathbf{v}, T(\mathbf{v})\}$ es linealmente dependiente si y solo si $T(\mathbf{v}) = \mathbf{v}$ o $T(\mathbf{v}) = \mathbf{0}$.

(**b**) De un ejemplo de tal transformación lineal con $V = \mathbb{R}^2$.

*El conjunto de todas las transformaciones lineales de un espacio vectorial V a un espacio vectorial W se denota mediante $\mathscr{L}(V, W)$. Si S y T están en $\mathscr{L}(V, W)$ puede definirse la **suma** S + T de S y T mediante*

$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$$

*para todo **v** en V. Si c es un escalar, se define el **múltiplo escalar** cT de T por c como*

$$(cT)(\mathbf{v}) = cT(\mathbf{v})$$

*para todo **v** en V. Entonces S + T y cT son ambas transformaciones de V a W.*

34. Demuestre que $S + T$ y cT son transformaciones lineales.

35. Demuestre que $\mathscr{L}(V, W)$ es un espacio vectorial con esta suma y multiplicación por un escalar.

36. Sean R, S y T transformaciones lineales tales que las siguientes operaciones tienen sentido. Demuestre que:

(**a**) $R \circ (S + T) = R \circ S + R \circ T$

(**b**) $c(R \circ S) = (cR) \circ S = R \circ (cS)$ para cualquier escalar c

6.5 El kernel y el rango de una transformación lineal

El espacio nulo y el espacio columna son dos de los subespacios fundamentales asociados con una matriz. En esta sección se extienden las nociones al kernel y el rango de una transformación lineal.

La palabra *kernel* (núcleo) se deriva de la palabra en inglés antiguo *cyrnel*, una forma de la palabra *corn*, que significa "semilla" o "grano". Como el núcleo del maíz, el kernel de una transformación lineal es su "núcleo" o "semilla" en el sentido de que aporta información acerca de muchas de las propiedades importantes de la transformación.

Definición Sea $T : V \to W$ una transformación lineal. El ***kernel*** de T, denotado $\ker(T)$, es el conjunto de todos los vectores en V que se mapean mediante T a $\mathbf{0}$ en W. Esto es,

$$\ker(T) = \{\mathbf{v} \text{ en } V : T(\mathbf{v}) = \mathbf{0}\}$$

El ***rango*** de T, denotado $\text{rango}(T)$, es el conjunto de todos los vectores en W que son imágenes de los vectores en V bajo T. Esto es,

$$\begin{aligned}\text{rango}(T) &= \{T(\mathbf{v}) : \mathbf{v} \text{ en } V\} \\ &= \{\mathbf{w} \text{ en } W : \mathbf{w} = T(\mathbf{v}) \text{ para algún } \mathbf{v} \text{ en } V\}\end{aligned}$$

Ejemplo 6.59

Sea A una matriz de $m \times n$ y sea $T = T_A$ la correspondiente transformación matricial de $\mathbb{R}^n$ a $\mathbb{R}^m$ definida mediante $T(\mathbf{v}) = A\mathbf{v}$. Entonces, como se vio en el capítulo 3, el rango de T es el espacio columna de A.

El kernel de T es

$$\begin{aligned}\ker(T) &= \{\mathbf{v} \text{ en } \mathbb{R}^n : T(\mathbf{v}) = \mathbf{0}\} \\ &= \{\mathbf{v} \text{ en } \mathbb{R}^n : A\mathbf{v} = \mathbf{0}\} \\ &= \text{nulo}(A)\end{aligned}$$

En palabras, el kernel de una transformación matricial es justo el espacio nulo de la matriz correspondiente.

Ejemplo 6.60

$\frac{dy}{dx}$

Encuentre el kernel y el rango del operador diferencial $D : \mathscr{P}_3 \to \mathscr{P}_2$ definido por $D(p(x)) = p'(x)$.

Solución Dado que $D(a + bx + cx^2 + dx^3) = b + 2cx + 3dx^2$, se tiene

$$\begin{aligned}\ker(D) &= \{a + bx + cx^2 + dx^3 : D(a + bx + cx^2 + dx^3) = 0\} \\ &= \{a + bx + cx^2 + dx^3 : b + 2cx + 3dx^2 = 0\}\end{aligned}$$

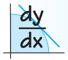
Pero $b + 2cx + 3dx^2 = 0$ si y sólo si $b = 2c = 3d = 0$, lo cual implica que $b = c = d = 0$. Por tanto,

$$\begin{aligned}\ker(D) &= \{a + bx + cx^2 + dx^3 : b = c = d = 0\} \\ &= \{a : a \text{ en } \mathbb{R}\}\end{aligned}$$

En otras palabras, el kernel de D es el conjunto de polinomios constantes.

El rango de D es todo de $\mathscr{P}_2$, pues *todo* polinomio en $\mathscr{P}_2$ es la imagen bajo D (esto es, la derivada) de *algún* polinomio en $\mathscr{P}_3$. Para ser específico, si $a + bx + cx^2$ está en $\mathscr{P}_2$, entonces

$$a + bx + cx^2 = D\left(ax + \left(\frac{b}{2}\right)x^2 + \left(\frac{c}{3}\right)x^3\right)$$



Ejemplo 6.61

Sea $S : \mathscr{P}_1 \to \mathbb{R}$ la transformación lineal definida por

$$S(p(x)) = \int_0^1 p(x)\,dx$$

Encuentre el kernel y el rango de S.

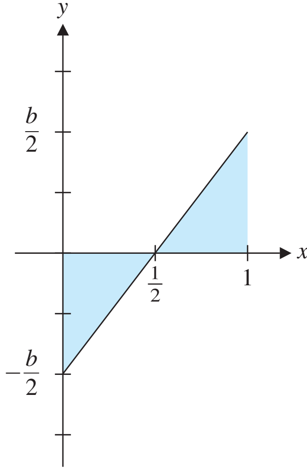


Figura 6.7

Si $y = -\frac{b}{2} + bx$,

entonces $\int_0^1 y\,dx = 0$

Solución En detalle, se tiene

$$\begin{aligned} S(a + bx) &= \int_0^1 (a + bx)\,dx \\ &= \left[ax + \frac{b}{2}x^2\right]_0^1 \\ &= \left(a + \frac{b}{2}\right) - 0 = a + \frac{b}{2} \end{aligned}$$

Por tanto,

$$\begin{aligned} \ker(S) &= \{a + bx : S(a + bx) = 0\} \\ &= \left\{a + bx : a + \frac{b}{2} = 0\right\} \\ &= \left\{a + bx : a = -\frac{b}{2}\right\} \\ &= \left\{-\frac{b}{2} + bx\right\} \end{aligned}$$

Geométricamente, $\ker(S)$ consiste de todos aquellos polinomios lineales cuyas gráficas tienen la propiedad de que el área entre la recta y el eje x está igualmente distribuida arriba y abajo del eje en el intervalo $[0, 1]$ (vea la figura 6.7).

El rango de S es $\mathbb{R}$, pues todo número real puede obtenerse como la imagen bajo S de algún polinomio en $\mathscr{P}_1$. Por ejemplo, si a es un número real arbitrario, entonces

$$\int_0^1 a\,dx = [ax]_0^1 = a - 0 = a$$

de modo que $a = S(a)$.

Ejemplo 6.62

Sea $T : M_{22} \to M_{22}$ la transformación lineal definida al tomar transpuestas: $T(A) = A^T$. Encuentre el kernel y el rango de T.

Solución Se ve que

$$\begin{aligned} \ker(T) &= \{A \text{ en } M_{22} : T(A) = O\} \\ &= \{A \text{ en } M_{22} : A^T = O\} \end{aligned}$$

Pero si $A^T = O$, entonces $A = (A^T)^T = O^T = O$. Se concluye que $\ker(T) = \{O\}$.

En consecuencia, para cualquier matriz A en M_{22}, se tiene $A = (A^T)^T = T(A^T)$ (y A^T está en M_{22}), se deduce que $\text{rango}(T) = M_{22}$.

En todos estos ejemplos, el kernel y el rango de una transformación lineal son subespacios del dominio y el codominio, respectivamente, de la transformación. Dado que se generaliza el espacio nulo y el espacio columna de una matriz, acaso esto no sea una sorpresa. No obstante, no debe dar nada por sentado, así que es necesario demostrar que no es una coincidencia.

Teorema 6.18 Sea $T : V \to W$ una transformación lineal. Entonces:

a. El kernel de T es un subespacio de V.
b. El rango de T es un subespacio de W.

Demostración (a) Dado que $T(\mathbf{0}) = \mathbf{0}$, el vector cero de V está en $\ker(T)$, de modo que $\ker(T)$ no está vacío. Sean $\mathbf{u}$ y $\mathbf{v}$ que están en $\ker(T)$ y sea c un escalar. Entonces $T(\mathbf{u}) = T(\mathbf{v}) = \mathbf{0}$, de modo que

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

y

$$T(c\mathbf{u}) = cT(\mathbf{u}) = c\mathbf{0} = \mathbf{0}$$

Por tanto, $\mathbf{u} + \mathbf{v}$ y $c\mathbf{u}$ están en $\ker(T)$ y $\ker(T)$ es un subespacio de V.

(b) Puesto que $\mathbf{0} = T(\mathbf{0})$, el vector cero de W está en rango(T), de modo que rango(T) no esta vacío. Sean $T(\mathbf{u})$ y $T(\mathbf{v})$ que están en el rango de T y sea c un escalar. Entonces $T(\mathbf{u}) + T(\mathbf{v}) = T(\mathbf{u} + \mathbf{v})$ es la imagen del vector $\mathbf{u} + \mathbf{v}$. Dado que $\mathbf{u}$ y $\mathbf{v}$ están en V, también lo está $\mathbf{u} + \mathbf{v}$ y por tanto $T(\mathbf{u}) + T(\mathbf{v})$ está en rango(T). De manera similar $cT(\mathbf{u}) = T(c\mathbf{u})$. Dado que $\mathbf{u}$ está en V, también lo está $c\mathbf{u}$, y en consecuencia $cT(\mathbf{u})$ está en rango(T). Por tanto, rango(T) es un subconjunto no vacío de W que está cerrado bajo la suma y la multiplicación por un escalar y por ende, es un subespacio de W.

La figura 6.8 ofrece una representación esquemática del kernel y el rango de una transformación lineal.

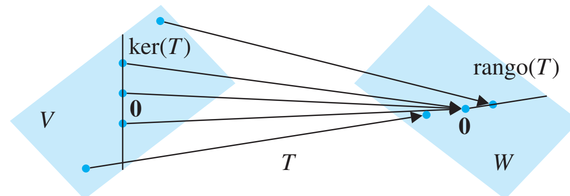


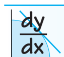
Figura 6.8
El kernel y el rango de $T : V \to W$

En el capítulo 3 se definió el rank de una matriz como la dimensión de su espacio columna y la nulidad de una matriz como la dimensión de su espacio nulo. Ahora se extienden dichas definiciones a transformaciones lineales.

Definición Sea $T : V \to W$ una transformación lineal. El ***rank*** de T es la dimensión del rango de T y se denota mediante rank(T). La ***nulidad*** de T es la dimensión del kernel de T y se denota mediante nulidad(T).

Ejemplo 6.63 Si A es una matriz y $T = T_A$ es la transformación matricial definida por $T(\mathbf{v}) = A\mathbf{v}$, entonces el rango y el kernel de T son el espacio columna y el espacio nulo de A, respectivamente, por el ejemplo 6.59. En consecuencia, de la sección 3.5, se tiene

$$\text{rank}(T) = \text{rank}(A) \quad \text{y} \quad \text{nulidad}(T) = \text{nulidad}(A)$$



Ejemplo 6.64

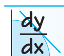
Encuentre el rank y la nulidad de la transformación lineal $D : \mathscr{P}_3 \to \mathscr{P}_2$ definida mediante $D(p(x)) = p'(x)$.

Solución En el ejemplo 6.60 calculó rango $(D) = \mathscr{P}_2$, de modo que

$$\text{rank}(D) = \dim \mathscr{P}_2 = 3$$

El kernel de D es el conjunto de todos los polinomios constantes: $\ker(D) = \{a : a \text{ en } \mathbb{R}\} = \{a \cdot 1 : a \text{ en } \mathbb{R}\}$. Por tanto, $\{1\}$ es una base para $\ker(D)$, de modo que

$$\text{nulidad(D)} = \dim(\ker(\text{D})) = 1$$



Ejemplo 6.65

Encuentre el rank y la nulidad de la transformación lineal $S : \mathscr{P}_1 \to \mathbb{R}$ definida por

$$S(p(x)) = \int_0^1 p(x)dx$$

Solución A partir del ejemplo 6.61, rango$(S) = \mathbb{R}$ y rank$(S) = \dim \mathbb{R} = 1$. Además,

$$\begin{aligned}\ker(S) &= \left\{-\frac{b}{2} + bx : b \text{ en } \mathbb{R}\right\} \\ &= \{b(-\tfrac{1}{2} + x) : b \text{ en } \mathbb{R}\} \\ &= \text{gen}(-\tfrac{1}{2} + x)\end{aligned}$$

de modo que $\{-\frac{1}{2} + x\}$ es una base para $\ker(S)$. Por tanto, nulidad$(S) = \dim(\ker(S)) = 1$.

Ejemplo 6.66

Encuentre el rank y la nulidad de la transformación lineal $T : M_{22} \to M_{22}$ definida mediante $T(A) = A^T$.

Solución En el ejemplo 6.62 se encontró que rango$(T) = M_{22}$ y $\ker(T) = \{O\}$. Por tanto,

$$\text{rank}(T) = \dim M_{22} = 4 \quad \text{y} \quad \text{nulidad}(T) = \dim\{O\} = 0$$

En el capítulo 3 se vio que el rank y la nulidad de una matriz A de $m \times n$ se relacionan mediante la fórmula rank(A) + nulidad$(A) = n$. Este es el Teorema del rank (Teorema 3.26). Dado que la transformación matricial $T = T_A$ tiene $\mathbb{R}^n$ como su dominio, podría rescribir la relación como

$$\text{rank}(A) + \text{nulidad}(A) = \dim \mathbb{R}^n$$

Esta versión del Teorema del rank se extiende de forma muy amigable a transformaciones lineales en general, como puede ver a partir de los últimos tres ejemplos:

$$\begin{aligned}&\text{rank}(D) + \text{nulidad}(D) = 3 + 1 = 4 = \dim \mathscr{P}_3 && \text{Ejemplo 6.64}\\ &\text{rank}(S) + \text{nulidad}(S) = 1 + 1 = 2 = \dim \mathscr{P}_1 && \text{Ejemplo 6.65}\\ &\text{rank}(T) + \text{nulidad}(T) = 4 + 0 = 4 = \dim \text{M}_{22} && \text{Ejemplo 6.66}\end{aligned}$$

Teorema 6.19 **Teorema del rank**

Sea $T: V \to W$ una transformación lineal de un espacio vectorial de dimensión finita V a un espacio vectorial W. Entonces

$$\text{rank}(T) + \text{nulidad}(T) = \dim V$$

En la siguiente sección verá cómo adaptar la demostración del Teorema 3.26 para probar esta versión del resultado. Por ahora, se ofrece una demostración alternativa que no usa matrices.

Demostración Sean $\dim V = n$ y $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ una base para $\ker(T)$ [de modo que nulidad$(T) = \dim(\ker(T)) = k$]. Puesto que $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es un conjunto linealmente independiente, puede extenderse a una base para V, por el Teorema 6.28. Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_n\}$ tal base. Si puede demostrar que el conjunto $\mathcal{C} = \{T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_n)\}$ es una base para rango(T), entonces se tendrá $\text{rank}(T) = \dim(\text{rango}(T)) = n - k$ y por tanto

$$\text{rank}(T) + \text{nulidad}(T) = k + (n - k) = n = \dim V$$

como se requiere.

Ciertamente, $\mathcal{C}$ está contenido en el rango de T. Para demostrar que $\mathcal{C}$ genera el rango de T, sea $T(\mathbf{v})$ un vector en el rango de T. Entonces $\mathbf{v}$ está en V, y dado que $\mathcal{B}$ es una base para V, es posible encontrar escalares $c_1, \ldots, c_n$ tales que

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n$$

Puesto que $\mathbf{v}_1, \ldots, \mathbf{v}_k$ están en el kernel de T, se tiene $T(\mathbf{v}_1) = \cdots = T(\mathbf{v}_k) = \mathbf{0}$, de modo que

$$\begin{aligned} T(\mathbf{v}) &= T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k + c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n) \\ &= c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k) + c_{k+1}T(\mathbf{v}_{k+1}) + \cdots + c_nT(\mathbf{v}_n) \\ &= c_{k+1}T(\mathbf{v}_{k+1}) + \cdots + c_nT(\mathbf{v}_n) \end{aligned}$$

Esto demuestra que $\mathcal{C}$ genera el rango de T.

Para demostrar que $\mathcal{C}$ es linealmente independiente, suponga que existen escalares $c_{k+1}, \ldots, c_n$ tales que

$$c_{k+1}T(\mathbf{v}_{k+1}) + \cdots + c_nT(\mathbf{v}_n) = \mathbf{0}$$

Entonces $T(c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n) = \mathbf{0}$, lo cual significa que $c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n$ está en el kernel de T y, en consecuencia, puede expresarse como una combinación lineal de los vectores base $\mathbf{v}_1, \ldots, \mathbf{v}_k$ de $\ker(T)$, por decir,

$$c_{k+1}\mathbf{v}_{k+1} + \cdots + c_n\mathbf{v}_n = c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$$

Pero ahora $$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k - c_{k+1}\mathbf{v}_{k+1} - \cdots - c_n\mathbf{v}_n = \mathbf{0}$$

y la independencia lineal de $\mathcal{B}$ fuerza $c_1 = \cdots = c_n = 0$. En particular, $c_{k+1} = \cdots = c_n = 0$, lo cual significa que $\mathcal{C}$ es linealmente independiente.

Se demostró que $\mathcal{C}$ es una base para el rango de T, de modo que, por los comentarios anteriores, la demostración está completa.

Se verificó el Teorema del rank para los ejemplos 6.64, 6.65 y 6.66. En la práctica, este teorema permite encontrar el rank y la nulidad de una transformación lineal sólo con la mitad del trabajo. Los siguientes ejemplos ilustran el proceso.

Ejemplo 6.67

Encuentre el rank y la nulidad de la transformación lineal $T : \mathscr{P}_2 \to \mathscr{P}_3$ definida mediante $T(p(x)) = xp(x)$. (Compruebe que T realmente es lineal.)

Solución En detalle, se tiene

$$T(a + bx + cx^2) = ax + bx^2 + cx^3$$

Se tiene que

$$\begin{aligned}\ker(T) &= \{a + bx + cx^2 : T(a + bx + cx^2) = 0\} \\ &= \{a + bx + cx^2 : ax + bx^2 + cx^3 = 0\} \\ &= \{a + bx + cx^2 : a = b = c = 0\} \\ &= \{0\}\end{aligned}$$

de modo que se tiene nulidad$(T) = \dim(\ker(T)) = 0$. El Teorema del rank implica que

$$\text{rank(T)} = \dim \mathscr{P}_2 - \text{nulidad(T)} = 3 - 0 = 3$$

Comentario En el ejemplo 6.67 sería igualmente sencillo encontrar primero el rank de T, pues $\{x, x^2, x^3\}$ es fácilmente vista como una base para el rango de T. Sin embargo, por lo general, uno de los dos (el rank o la nulidad de una transformación lineal) sería más sencillo de calcular; entonces puede usar el Teorema del rank para encontrar el otro. Con práctica, mejorará para saber por cuál camino avanzar.

Ejemplo 6.68

Sea W el espacio vectorial de todas las matrices simétricas de 2×2. Defina una transformación lineal $T : W \to \mathscr{P}_2$ mediante

$$T\begin{bmatrix} a & b \\ b & c \end{bmatrix} = (a - b) + (b - c)x + (c - a)x^2$$

(Compruebe que T es lineal.) Encuentre el rank y la nulidad de T.

Solución La nulidad de T es más fácil de calcular directamente que el rank, así que se procede del modo siguiente:

$$\begin{aligned}\ker(T) &= \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : T\begin{bmatrix} a & b \\ b & c \end{bmatrix} = 0\right\} \\ &= \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : (a - b) + (b - c)x + (c - a)x^2 = 0\right\} \\ &= \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : (a - b) = (b - c) = (c - a) = 0\right\} \\ &= \left\{\begin{bmatrix} a & b \\ b & c \end{bmatrix} : a = b = c\right\} \\ &= \left\{\begin{bmatrix} c & c \\ c & c \end{bmatrix}\right\} = \text{gen}\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)\end{aligned}$$

Por tanto, $\left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$ es una base para el kernel de T, de modo que nulidad$(T) = \dim(\ker(T)) = 1$. El Teorema del rank y el ejemplo 6.42 dicen que rank$(T) = \dim W -$ nulidad$(T) = 3 - 1 = 2$.

Transformaciones lineales inyectivas y sobreyectivas

Ahora se investigarán criterios para que una transformación lineal sea invertible. Las claves para la discusión son las muy importantes propiedades inyectivas y sobreyectivas.

Definición Una transformación lineal $T: V \rightarrow W$ se llama ***inyectiva*** si T mapea distintos vectores en V a distintos vectores en W. Si rango$(T) = W$, entonces T se llama ***sobreyectiva***.

Comentarios

- La definición de inyectiva puede escribirse de manera más formal del modo siguiente:

$T: V \rightarrow W$ es inyectiva si, para todos $\mathbf{u}$ y $\mathbf{v}$ en V,

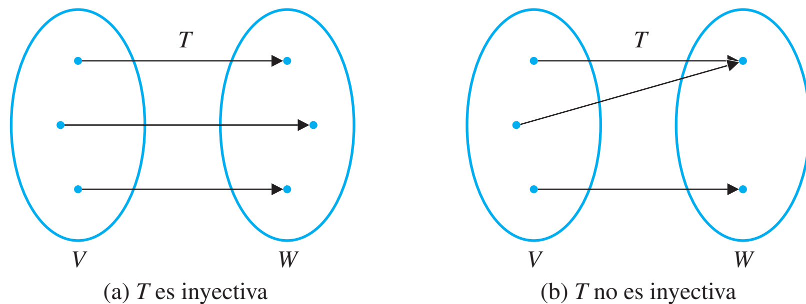
$$\mathrm{u} \neq \mathbf{v} \text{ implica que } T(\mathbf{u}) \neq T(\mathbf{v})$$

El enunciado anterior es equivalente al siguiente:

$T: V \rightarrow W$ es inyectiva si, para todos $\mathbf{u}$ y $\mathbf{v}$ en V,

$$T(\mathrm{u}) = (\mathbf{v}) \text{ implica que } \mathbf{u} = \mathbf{v}$$

La figura 6.9 ilustra estos dos enunciados.



(a) T es inyectiva (b) T no es inyectiva

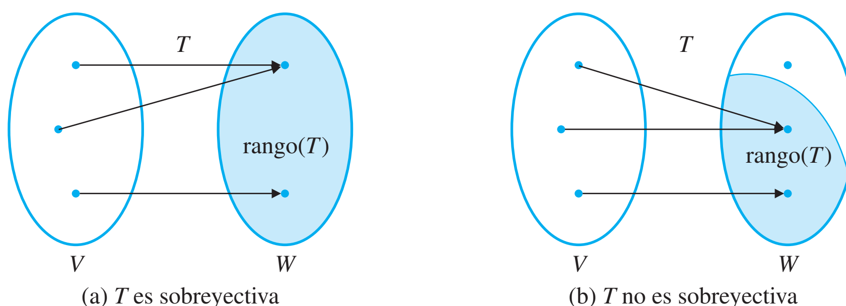
Figura 6.9

- Otra forma de escribir la definición de sobreyectiva es la siguiente:

$T: V \rightarrow W$ es sobreyectiva si, para todo $\mathbf{w}$ en W, existe al menos un $\mathbf{v}$ en V tal que

$$\mathbf{w} = T(\mathbf{v})$$

En otras palabras, *dado* $\mathbf{w}$ en W, ¿existe algún $\mathbf{v}$ en V tal que $\mathbf{w} = T(\mathbf{v})$? Si, para un $\mathbf{w}$ arbitrario, puede resolver esta ecuación para $\mathbf{v}$, entonces T es sobreyectiva (vea la figura 6.10).



(a) T es sobreyectiva (b) T no es sobreyectiva

Figura 6.10

Ejemplo 6.69

¿Cuál de las siguientes transformaciones lineales son inyectivas? ¿y sobreyectivas?

(a) $T: \mathbb{R}^2 \to \mathbb{R}^3$ definida mediante $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x \\ x - y \\ 0 \end{bmatrix}$

(b) $D: \mathcal{P}_3 \to \mathcal{P}_2$ definida mediante $D(p(x)) = p'(x)$

(c) $T: M_{22} \to M_{22}$ definida mediante $T(A) = A^T$

Solución (a) Sea $T\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = T\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$. Entonces

$$\begin{bmatrix} 2x_1 \\ x_1 - y_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2x_2 \\ x_2 - y_2 \\ 0 \end{bmatrix}$$

de modo que $2x_1 = 2x_2$ y $x_1 - y_1 = x_2 - y_2$. Al resolver estas ecuaciones se ve que $x_1 = x_2$ y $y_1 = y_2$. Por tanto, $\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}$ y en consecuencia T es inyectiva.

T no es sobreyectiva, pues su rango no es todo $\mathbb{R}^3$. Para ser específicos, no hay vector $\begin{bmatrix} x \\ y \end{bmatrix}$ en $\mathbb{R}^2$ tal que $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. (¿Por qué no?)

(b) En el ejemplo 6.60 se demostró que rango$(D) = \mathcal{P}_2$, de modo que D es sobreyectiva. D no es inyectiva pues distintos polinomios en $\mathcal{P}_3$ pueden tener la misma derivada. Por ejemplo, $x^3 \neq x^3 + 1$, pero $D(x^3) = 3x^2 = D(x^3 + 1)$.

(c) Sean A y B en M_{22}, con $T(A) = T(B)$. Entonces $A^T = B^T$, de modo que $A = (A^T)^T = (B^T)^T = B$. En consecuencia, T es inyectiva. En el ejemplo 6.62 se demostró que rango$(T) = M_{22}$. Por tanto, T es sobreyectiva.

Es evidente que existe un criterio muy simple para determinar si una transformación lineal es inyectiva.

Teorema 6.20 Una transformación lineal $T: V \to W$ es inyectiva si y sólo si $\ker(T) = \{\mathbf{0}\}$.

Demostración Suponga que T es inyectiva. Si $\mathbf{v}$ está en el kernel de T, entonces $T(\mathbf{v}) = \mathbf{0}$. Pero también se sabe que $T(\mathbf{0}) = \mathbf{0}$, de modo que $T(\mathbf{v}) = T(\mathbf{0})$. Puesto que T es inyectiva, esto implica que $\mathbf{v} = \mathbf{0}$, de modo que el único vector en el kernel de T es el vector cero.

Por el contrario, suponga que $\ker(T) = \{\mathbf{0}\}$. Para demostrar que T es inyectiva, sean $\mathbf{u}$ y $\mathbf{v}$ que están en V, con $T(\mathbf{u}) = T(\mathbf{v})$. Entonces $T(\mathbf{u} - \mathbf{v}) = T(\mathbf{u}) - T(\mathbf{v}) = \mathbf{0}$, lo que implica que $\mathbf{u} - \mathbf{v}$ está en el kernel de T. Pero $\ker(T) = \{\mathbf{0}\}$, de modo que debe tener $\mathbf{u} - \mathbf{v} = \mathbf{0}$ o, de manera equivalente, $\mathbf{u} = \mathbf{v}$. Esto demuestra que T es inyectiva.

Ejemplo 6.70

Demuestre que la transformación lineal $T : \mathbb{R}^2 \to \mathscr{P}_1$ definida mediante

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x$$

es inyectiva y sobreyectiva.

Solución Si $\begin{bmatrix} a \\ b \end{bmatrix}$ está en el kernel de T, entonces

$$0 = T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x$$

Se tiene que $a = 0$ y $a + b = 0$. Por tanto, $b = 0$ y en consecuencia $\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Así que, $\ker(T) = \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$, y T es inyectiva por el Teorema 6.20.

Por el Teorema del rank,

$$\text{rank}(T) = \dim \mathbb{R}^2 - \text{nulidad}(T) = 2 - 0 = 2$$

Por tanto, el rango de T es un subespacio bidimensional de $\mathbb{R}^2$ y en consecuencia $\text{rango}(T) = \mathbb{R}^2$. Se tiene que T es sobreyectiva.

Para transformaciones lineales entre dos espacios vectoriales n-dimensionales, las propiedades de inyectivas y sobreyectivas están estrechamente relacionadas. Observe primero que para una transformación lineal $T : V \to W$, $\ker(T) = \{\mathbf{0}\}$ si y sólo si $\text{nulidad}(T) = 0$, y T es sobreyectiva si y sólo si $\text{rank}(T) = \dim W$. (¿Por qué?) La demostración del siguiente teorema en esencia usa el método del ejemplo 6.70.

Teorema 6.21

Sea $\dim V = \dim W = n$. Entonces una transformación lineal $T : V \to W$ es inyectiva si y sólo si es sobreyectiva.

Demostración Suponga que T es inyectiva. Entonces $\text{nulidad}(T) = 0$ por el Teorema 6.20 y la observación anterior al teorema 6.21. El Teorema del rank implica que

$$\text{rank}(T) = \dim V - \text{nulidad}(T) = n - 0 = n$$

Por tanto, T es sobreyectiva.

Por el contrario, suponga que T es sobreyectiva. Entonces $\text{rank}(T) = \dim W = n$. Por el Teorema del rank,

$$\text{nulidad}(T) = \dim V - \text{rank}(T) = n - n = 0$$

En consecuencia, $\ker(T) = \{\mathbf{0}\}$ y T es inyectiva.

En la sección 6.4 se apuntó que si $T : V \to W$ es una transformación lineal, entonces la imagen de una base para V bajo T no necesita ser una base para el rango de T. Ahora es posible dar una condición que garantice que una base para V se mapeará mediante T a una base para W.

Teorema 6.22

Sea $T : V \to W$ una transformación lineal inyectiva. Si $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ es un conjunto linealmente independiente en V, entonces $T(S) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ es un conjunto linealmente independiente en W.

Demostración Sean $c_1, \ldots, c_k$ escalares tales que

$$c_1T(\mathbf{v}_1) + \cdots + c_kT(\mathbf{v}_k) = \mathbf{0}$$

Entonces $T(c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k) = \mathbf{0}$, lo que implica que $c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k$ está en el kernel de T. Pero, dado que T es inyectiva, ker $(T) = \{\mathbf{0}\}$, por el Teorema 6.20. En consecuencia,

$$c_1\mathbf{v}_1 + \cdots + c_k\mathbf{v}_k = \mathbf{0}$$

Pero, puesto que dim $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. es linealmente independiente, todos los escalares c_i deben ser 0. Por tanto, $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_k)\}$ es linealmente independiente.

Corolario 6.23

Sea dim V = dim $W = n$. Entonces una transformación lineal inyectiva $T : V \to W$ mapea una base para V a una base para W.

Demostración Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base para V. Por el Teorema 6.22, $T(\mathcal{B}) = \{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ es un conjunto linealmente independiente en W, así que sólo es necesario demostrar que $T(\mathcal{B})$ genera a W. Pero, por el Teorema 6.15, $T(\mathcal{B})$ genera el rango de T. Más aún, T es sobreyectiva, por el Teorema 6.21, de modo que rango$(T) = W$. Por tanto, $T(\mathcal{B})$ genera a W, lo que completa la demostración.

Ejemplo 6.71

Sea $T : \mathbb{R}^2 \to \mathscr{P}_1$ una transformación lineal del ejemplo 6.70, definida mediante

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x$$

Entonces, por el corolario 6.23, la base estándar $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2\}$ para $\mathbb{R}^2$ se mapea a una base $T(\mathcal{E}) = \{T(\mathbf{e}_1), T(\mathbf{e}_2)\}$ de $\mathscr{P}_1$. Se descubre que

$$T(\mathbf{e}_1) = T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 + x \quad \text{y} \quad T(\mathbf{e}_2) = T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = x$$

Se concluye que $\{1 + x, x\}$ es una base para $\mathscr{P}_1$.

Ahora puede determinar cuáles transformaciones lineales $T : V \to W$ son invertibles.

Teorema 6.24

Una transformación lineal $T : V \to W$ es invertible si y sólo si es inyectiva y sobreyectiva.

Demostración Suponga que T es invertible. Entonces existe una transformación lineal $T^{-1}: W \rightarrow V$ tal que

$$T^{-1} \circ T = I_V \quad \text{y} \quad T \circ T^{-1} = I_W$$

Para demostrar que T es inyectiva, sea $\mathbf{v}$ que está en el kernel de T. Entonces, $T(\mathbf{v}) = \mathbf{0}$. Por tanto,

$$\begin{aligned} T^{-1}(T(\mathbf{v})) = T^{-1}(\mathbf{0}) &\Rightarrow (T^{-1} \circ T)(\mathbf{v}) = \mathbf{0} \\ &\Rightarrow I(\mathbf{v}) = \mathbf{0} \\ &\Rightarrow \mathbf{v} = \mathbf{0} \end{aligned}$$

que establece que $\ker(T) = \{\mathbf{0}\}$. En consecuencia, T es inyectiva por el Teorema 6.20.

Para demostrar que T es sobreyectiva, sea $\mathbf{w}$ que está en W y sea $\mathbf{v} = T^{-1}(\mathbf{w})$. Entonces

$$\begin{aligned} T(\mathbf{v}) &= T(T^{-1}(\mathbf{w})) \\ &= (T \circ T^{-1})(\mathbf{w}) \\ &= I(\mathbf{w}) \\ &= \mathbf{w} \end{aligned}$$

lo que demuestra que $\mathbf{w}$ es la imagen de $\mathbf{v}$ bajo T. Puesto que $\mathbf{v}$ está en V, esto demuestra que T es sobreyectiva.

Por el contrario, suponga que T es inyectiva y sobreyectiva. Esto significa que nulidad$(T) = 0$ y rank$(T) = \dim W$. Es necesario demostrar que existe una transformación lineal $T': W \rightarrow V$ tal que $T' \circ T = I_V$ y $T \circ T' = I_W$.

Sea $\mathbf{w}$ que está en W. Dado que T es sobreyectiva, existe algún vector $\mathbf{v}$ en V tal que $T(\mathbf{v}) = \mathbf{w}$. Sólo existe uno de tales vectores $\mathbf{v}$, pues, si $\mathbf{v}'$ es otro vector en V tal que $T(\mathbf{v}') = \mathbf{w}$, entonces $T(\mathbf{v}) = T(\mathbf{v}')$; el hecho de que T sea inyectiva implica entonces que $\mathbf{v} = \mathbf{v}'$. En consecuencia, tiene sentido definir un mapeo $T': W \rightarrow V$ al establecer $T'(\mathbf{w}) = \mathbf{v}$.

Se tiene que

$$(T' \circ T)(\mathbf{v}) = T'(T(\mathbf{v})) = T'(\mathbf{w}) = \mathbf{v}$$

y

$$(T \circ T')(\mathbf{w}) = T(T'(\mathbf{w})) = T(\mathbf{v}) = \mathbf{w}$$

Entonces se tiene que $T' \circ T = I_V$ y $T \circ T' = I_W$. Ahora debe demostrar que T' es una transformación *lineal*.

Para este fin, sean $\mathbf{w}_1$ y $\mathbf{w}_2$ que están en W y sean c_1 y c_2 escalares. Como anteriormente, sean $T(\mathbf{v}_1) = \mathbf{w}_1$ y $T(\mathbf{v}_2) = \mathbf{w}_2$. Entonces $\mathbf{v}_1 = T'(\mathbf{w}_1)$ y $\mathbf{v}_2 = T'(\mathbf{w}_2)$ y

$$\begin{aligned} T'(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) &= T'(c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2)) \\ &= T'(T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)) \\ &= I(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) \\ &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \\ &= c_1T'(\mathbf{w}_1) + c_2T'(\mathbf{w}_2) \end{aligned}$$

En consecuencia, T' es lineal y, por el Teorema 6.17, $T' = T^{-1}$.

Las palabras *isomorfismo* e *isomórfico* provienen de las palabras griegas *isos*, que significa "igual", y *morph*, que significa "formar". Por tanto, en sentido figurado, los espacios vectoriales isomórficos tienen "formas iguales".

Isomorfismos de espacios vectoriales

Ahora está en posibilidad de describir, en términos concretos, qué significa que dos espacios vectoriales sean "esencialmente iguales".

Definición Una transformación lineal $T : V \to W$ se llama ***isomorfismo*** si es inyectiva y sobreyectiva. Si V y W son dos espacios vectoriales tales que existe un isomorfismo de V a W, entonces se dice que V es ***isomórfica*** a W y se escribe $V \cong W$.

Ejemplo 6.72

Demuestre que $\mathscr{P}_{n-1}$ y $\mathbb{R}^n$ son isomórficas.

Solución El proceso de formar el vector coordenado de un polinomio proporciona un posible isomorfismo (como ya se observó en la sección 6.2, aunque ahí no se usó el término *isomorfismo*). Específicamente, defina $T : \mathscr{P}_{n-1} \to \mathbb{R}^n$ mediante $T(p(x)) = [p(x)]_{\mathcal{E}}$, donde $\mathcal{E} = \{1, x, \ldots, x^{n-1}\}$ es la base estándar para $\mathscr{P}_{n-1}$. Esto es,

$$T(a_0 + a_1x + \cdots + a_{n-1}x^{n-1}) = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix}$$

El Teorema 6.6 muestra que T es una transformación lineal. Si $p(x) = a_0 + a_1x + \cdots + a_{n-1}x^{n-1}$ está en el kernel de T, entonces

$$\begin{bmatrix} a_0 \\ \vdots \\ a_{n-1} \end{bmatrix} = T(a_0 + a_1x + \cdots + a_{n-1}x^{n-1}) = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Por ende, $a_0 = a_1 = \cdots = a_{n-1} = 0$, de modo que $p(x) = 0$. En consecuencia, $\ker(T) = \{0\}$ y T es inyectiva. Puesto que $\dim \mathscr{P}_{n-1} = \dim \mathbb{R}^n = n$, T también es sobreyectiva, por el Teorema 6.21. Por tanto, T es un isomorfismo y $\mathscr{P}_{n-1} \cong \mathbb{R}^n$.

Ejemplo 6.73

Demuestre que M_{mn} y $\mathbb{R}^{mn}$ son isomórficos.

Solución Una vez más, el mapeo coordenado desde M_{mn} hacia $\mathbb{R}^{mn}$ (como en el ejemplo 6.36) es un isomorfismo. Los detalles de la demostración se dejan como ejercicio.

De hecho, la forma más sencilla de decir si dos espacios vectoriales son isomórficos es simplemente verificar sus dimensiones, como muestra el siguiente teorema.

Teorema 6.25 Sean V y W dos espacios vectoriales de dimensión finita (sobre el mismo campo de escalares). Entonces V es isomórfico a W si y sólo si $\dim V = \dim W$.

Demostración Sea $n = \dim V$. Si V es isomórfico a W, entonces existe un isomorfismo $T: V \to W$. Dado que T es inyectiva, nulidad$(T) = 0$. Entonces el Teorema del rank implica que

$$\text{rank}(T) = \dim V - \text{nulidad}(T) = n - 0 = n$$

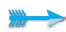
Por tanto, el rango de T es un subespacio n-dimensional de W. Pero, puesto que T es sobreyectiva, $W = \text{rango}(T)$, así que $\dim W = n$, como se quería demostrar.

Por el contrario, suponga que V y W tienen la misma dimensión, n. Sea $\mathcal{B} = \{\mathbf{v}1, \ldots, \mathbf{v}_n\}$ una base para V y sea $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ una base para W. Se definirá una transformación lineal $T: V \to W$ y entonces se demostrará que T es inyectiva y sobreyectiva. Un vector arbitrario $\mathbf{v}$ en V puede escribirse de manera única como una combinación lineal de los vectores en la base $\mathcal{B}$, por decir,

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n$$

Se define T mediante

$$T(\mathbf{v}) = c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n$$



Es directo comprobar que T es lineal. (Hágalo.) Para ver que T es inyectiva, suponga que $\mathbf{v}$ está en el kernel de T. Entonces

$$c_1\mathbf{w}_1 + \cdots + c_n\mathbf{w}_n = T(\mathbf{v}) = \mathbf{0}$$

y la independencia lineal de $\mathcal{C}$ fuerza $c_1 = \cdots = c_n = 0$. Pero entonces

$$\mathbf{v} = c_1\mathbf{v}_1 + \cdots + c_n\mathbf{v}_n = \mathbf{0}$$

de modo que $\ker(T) = \{\mathbf{0}\}$, lo que significa que T es inyectiva. Puesto que $\dim V = \dim W$, T también es sobreyectiva, por el Teorema 6.21. Por tanto, T es un isomorfismo y $V \cong W$.

Ejemplo 6.74 Demuestre que $\mathbb{R}^n$ y $\mathcal{P}_n$ no son isomórficos.

Solución Dado que $\dim \mathbb{R}^n = n \neq n + 1 = \dim \mathcal{P}_n$, $\mathbb{R}^n$ y $\mathcal{P}_n$ no son isomórficos, por el Teorema 6.25.

Ejemplo 6.75 Sea W el espacio vectorial de todas las matrices simétricas de 2×2. Demuestre que W es isomórfico a $\mathbb{R}^3$.

Solución En el ejemplo 6.42 se demostró que $W = 3$. Por tanto, $\dim W = \dim \mathbb{R}^3$, de modo que $W \cong \mathbb{R}^3$, por el Teorema 6.25. (Existe un obvio candidato para un isomorfismo $T: W \to \mathbb{R}^3$. ¿Cuál es?)

Comentario Todos los ejemplos han sido espacios vectoriales *reales*, pero los teoremas que se demostraron son verdaderos para espacios vectoriales sobre los números complejos $\mathbb{C}$ o $\mathbb{Z}_p$, donde p es primo. Por ejemplo, el espacio vectorial $M_{22}(\mathbb{Z}_2)$ de todas las matrices de 2×2 con entradas de $\mathbb{Z}_2$ tiene dimensión 4 como espacio vectorial sobre $\mathbb{Z}_2$, y por tanto $M_{22}(\mathbb{Z}_2) \cong \mathbb{Z}_2^4$.

Ejercicios 6.5

1. Sea $T: M_{22} \to M_{22}$ la transformación lineal definida por

$$T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}$$

(a) ¿Cuáles, si hay alguna, de las siguientes matrices están en $\ker(T)$?

(i) $\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 0 & 4 \\ 2 & 0 \end{bmatrix}$ **(iii)** $\begin{bmatrix} 3 & 0 \\ 0 & -3 \end{bmatrix}$

(b) ¿Cuáles, si hay alguna, de las matrices en el inciso (a) están en rango(T)?

(c) Describa $\ker(T)$ y rango(T).

2. Sea $T: M_{22} \to \mathbb{R}$ definida por $T(A) = \text{tr}(A)$.

(a) ¿Cuáles, si hay alguna, de las siguientes matrices están en $\ker(T)$?

(i) $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 3 & -2 \\ 2 & 1 \end{bmatrix}$ **(iii)** $\begin{bmatrix} -1 & 2 \\ 0 & 1 \end{bmatrix}$

(b) ¿Cuáles, si hay alguno, de los siguientes escalares están en rango(T)?

(i) 0 **(ii)** 5 **(iii)** $-\sqrt{2}$

(c) Describa $\ker(T)$ y rango(T).

3. Sea $T: \mathscr{P}_2 \to \mathbb{R}^2$ la transformación lineal definida por

$$T(a + bx + cx^2) = \begin{bmatrix} a - b \\ b + c \end{bmatrix}$$

(a) ¿Cuáles, si hay alguno, de los siguientes polinomios están en $\ker(T)$?

(i) $1 + x$ **(ii)** $x - x^2$ **(iii)** $1 + x - x^2$

(b) ¿Cuáles, si hay alguno, de los siguientes vectores están en rango(T)?

(i) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ **(ii)** $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ **(iii)** $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

(c) Describa $\ker(T)$ y rango(T).

4. Sea $T: \mathscr{P}_2 \to \mathscr{P}_2$ la transformación lineal definida por $T(p(x)) = xp'(x)$.

(a) ¿Cuáles, si hay alguno, de los siguientes polinomios están en $\ker(T)$?

(i) 2 **(ii)** x^2 **(iii)** $1 - x$

(b) ¿Cuáles, si hay alguno, de los polinomios del inciso (a) están en rango(T)?

(c) Describa $\ker(T)$ y rango(T).

En los ejercicios 5-8, encuentre bases para el kernel y el rango de la transformación lineal T en los ejercicios indicados. En cada caso, establezca la nulidad y el rank de T y verifique el Teorema del rank.

5. Ejercicio 1 **6.** Ejercicio 2

7. Ejercicio 3 **8.** Ejercicio 4

En los ejercicios 9-14, encuentre o la nulidad o el rank de T y luego use el Teorema del rank para encontrar el otro.

9. $T: M_{22} \to \mathbb{R}^2$ definida mediante $T\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a - b \\ c - d \end{bmatrix}$

10. $T: \mathscr{P}_2 \to \mathbb{R}^2$ definida mediante $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$

11. $T: M_{22} \to M_{22}$ definida mediante $T(A) = AB$, donde $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

12. $T: M_{22} \to M_{22}$ definida mediante $T(A) = AB - BA$, donde $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$

13. $T: \mathscr{P}_2 \to \mathbb{R}$ definida mediante $T(p(x)) = p'(0)$

14. $T: M_{33} \to M_{33}$ definida mediante $T(A) = A - A^T$

En los ejercicios 15-20, determine si la transformación lineal T es (a) inyectiva y (b) sobreyectiva.

15. $T: \mathbb{R}^2 \to \mathbb{R}^2$ definida mediante $T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - y \\ x + 2y \end{bmatrix}$

16. $T : \mathscr{P}_2 \to M_{22}$ definida mediante

$$T(a + bx + cx^2) = \begin{bmatrix} a + b & a + 2c \\ 2a + c & b - c \end{bmatrix}$$

17. $T : \mathscr{P}_2 \to \mathbb{R}^3$ definida mediante

$$T(a + bx + cx^2) = \begin{bmatrix} 2a - b \\ a + b - 3c \\ c - a \end{bmatrix}$$

18. $T : \mathscr{P}_2 \to \mathbb{R}^2$ definida mediante $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$

19. $T : \mathbb{R}^3 \to M_{22}$ definida mediante

$$T\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} a - b & b - c \\ a + b & b + c \end{bmatrix}$$

20. $T : \mathbb{R}^3 \to W$ definida mediante $T\begin{bmatrix} a \\ b \\ c \end{bmatrix} =$

$\begin{bmatrix} a + b + c & b - 2c \\ b - 2c & a - c \end{bmatrix}$, donde W es el espacio vectorial de todas las matrices simétricas de 2×2

En los ejercicios 21-26, determine si V y W son isomórficos. Si lo son, ofrezca un isomorfismo $T : V \to W$ explícito.

21. $V = D_3$ (matrices diagonales de 3×3), $W = \mathbb{R}^3$

22. $V = S_3$ (matrices simétricas de 3×3), $W = U_3$ (matrices triangulares superiores de 3×3)

23. $V = S_3$ (matrices simétricas de 3×3), $W = S_3'$ (matrices antisimétricas de 3×3).

24. $V = \mathscr{P}_2$, $W = \{p(x) \text{ en } \mathscr{P}_3 : p(0) = 0\}$

a + bi **25.** $V = \mathbb{C}$, $W = \mathbb{R}^2$

26. $V = \{A \text{ en } M_{22} : \text{tr}(A) = 0\}$, $W = \mathbb{R}^2$

dy/dx **27.** Demuestre que $T : \mathscr{P}_n \to \mathscr{P}_n$ definida mediante $T(p(x)) = p(x) + p'(x)$ es un isomorfismo.

28. Demuestre que $T : \mathscr{P}_n \to \mathscr{P}_n$ definida mediante $T(p(x)) = p(x - 2)$ es un isomorfismo.

29. Demuestre que $T : \mathscr{P}_n \to \mathscr{P}_n$ definida mediante

$T(p(x)) = x^n p\left(\frac{1}{x}\right)$ es un isomorfismo.

30. **(a)** Demuestre que $\mathscr{C}[0, 1] \cong \mathscr{C}[2, 3]$. [*Sugerencia:* defina $T : \mathscr{C}[0, 1] \to \mathscr{C}[2, 3]$ al hacer $T(f)$ la función cuyo valor en x es $(T(f))(x) = f(x - 2)$ para x en $[2, 3]$.]

(b) Demuestre que $\mathscr{C}[0, 1] \cong \mathscr{C}[a, a + 1]$ para toda a.

31. Demuestre que $\mathscr{C}[0, 1] \cong \mathscr{C}[0, 2]$.

32. Demuestre que $\mathscr{C}[a, b] \cong \mathscr{C}[c, d]$ para toda $a < b$ y $c < d$.

33. Sean $S : V \to W$ y $T : U \to V$ transformaciones lineales.

(a) Demuestre que si S y T son ambas inyectivas, también lo es $S \circ T$.

(b) Demuestre que si S y T son ambas sobreyectivas, también lo es $S \circ T$.

34. Sean $S : V \to W$ y $T : U \to V$ transformaciones lineales.

(a) Demuestre que si $S \circ T$ es inyectiva, también lo es T.

(b) Demuestre que si $S \circ T$ es sobreyectiva, también lo es S.

35. Sea $T : V \to W$ una transformación lineal entre dos espacios vectoriales de dimensión finita.

(a) Demuestre que si $\dim V < \dim W$, entonces T no puede ser sobreyectiva.

(b) Demuestre que si $\dim V > \dim W$, entonces T no puede ser inyectiva.

36. Sean $a_0, a_1, \ldots, a_n$, $n + 1$ números reales distintos. Defina $T : \mathscr{P}_n \to \mathbb{R}^{n+1}$ mediante

$$T(p(x)) = \begin{bmatrix} p(a_0) \\ p(a_1) \\ \vdots \\ p(a_n) \end{bmatrix}$$

Demuestre que T es un isomorfismo.

37. Si V es un espacio vectorial de dimensión finita y $T : V \to V$ es una transformación lineal tal que $\text{rank}(T) = \text{rank}(T^2)$, Demuestre que $\text{rango}(T) \cap \ker(T) = \{\mathbf{0}\}$. [*Sugerencia:* T^2 denota $T \circ T$. Use el Teorema del rank para ayudar a demostrar que los kerneles de T y T^2 son iguales.]

38. Sean U y W subespacios de un espacio vectorial V de dimensión finita. Defina $T : U \times W \to V$ mediante $T(\mathbf{u}, \mathbf{w}) = \mathbf{u} - \mathbf{w}$.

(a) Demuestre que T es una transformación lineal.

(b) Demuestre que $\text{rango}(T) = U + W$.

(c) Demuestre que $\ker(T) \cong U \cap W$. [*Sugerencia:* vea el ejercicio 50 de la sección 6.1.]

(d) Demuestre la ***identidad de Grassmann***:

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W)$$

[*Sugerencia*: aplique el Teorema del rank y use los resultados de (a) y (b) y el ejercicio 43(b) de la sección 6.2.]

6.6 La matriz de una transformación lineal

El Teorema 6.15 demostró que una transformación lineal $T: V \to W$ está completamente determinada por su efecto sobre un conjunto generador para V. En particular, si se sabe cómo actúa T sobre una base para V, entonces es posible calcular $T(\mathbf{v})$ para cualquier vector $\mathbf{v}$ en V. El ejemplo 6.55 ilustró el proceso. En el Teorema 3.31 se usó de manera implícita esta importante propiedad de las transformaciones lineales para ayudar a calcular la matriz estándar de una transformación lineal $T: \mathbb{R}^n \to \mathbb{R}^m$. En esta sección se demostrará que toda transformación lineal entre espacios vectoriales de dimensión finita puede representarse como una transformación matricial.

Suponga que V es un espacio vectorial n-dimensional, W es un espacio vectorial m-dimensional y $T: V \to W$ es una transformación lineal. Sean $\mathcal{B}$ y $\mathcal{C}$ bases para V y W, respectivamente. Entonces el vector coordenado que mapea $R(\mathbf{v}) = [\mathbf{v}]_{\mathcal{B}}$ define un isomorfismo $R: V \to \mathbb{R}^n$. Al mismo tiempo, se tiene un isomorfismo $S: W \to \mathbb{R}^m$ dado por $S(\mathbf{w}) = [\mathbf{w}]_{\mathcal{C}}$, que permite asociar la imagen $T(\mathbf{v})$ con el vector $[T(\mathbf{v})]_{\mathcal{C}}$ en $\mathbb{R}^m$. La figura 6.11 ilustra las relaciones.

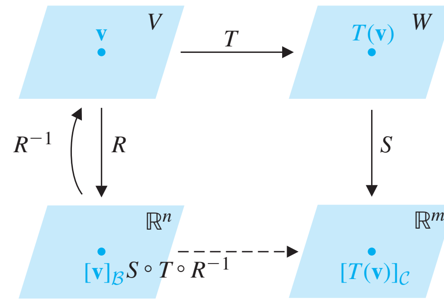


Figura 6.11

Dado que R es un isomorfismo, es invertible, de modo que puede formarse el mapeo compuesto

$$S \circ T \circ R^{-1}: \mathbb{R}^n \to \mathbb{R}^m$$

que mapea $[\mathbf{v}]_{\mathcal{B}}$ a $[T(\mathbf{v})]_{\mathcal{C}}$. Puesto que este mapeo va de $\mathbb{R}^n$ a $\mathbb{R}^m$, se sabe del capítulo 3 que es una transformación matricial. Entonces, ¿cuál es la matriz estándar de $S \circ T \circ R^{-1}$? Se quiere encontrar la matriz A de $m \times n$ tal que $A[\mathbf{v}]_{\mathcal{B}} = (S \circ T \circ R^{-1})([\mathbf{v}]_{\mathcal{B}})$. O, puesto que $(S \circ T \circ R^{-1})([\mathbf{v}]_{\mathcal{B}}) = [T(\mathbf{v})]_{\mathcal{C}}$, se requiere

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

Es evidente que es sorprendentemente fácil de encontrar. La idea básica es la del Teorema 3.31. Las columnas de A son las imágenes de los vectores base estándar para $\mathbb{R}^n$ bajo $S \circ T \circ R^{-1}$. Pero, si $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ es una base para V, entonces

$$\begin{aligned} R(\mathbf{v}_i) &= [\mathbf{v}_i]_{\mathcal{B}} \\ &= \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \quad \leftarrow i\text{-ésima entrada} \\ &= \mathbf{e}_i \end{aligned}$$

de modo que $R^{-1}(\mathbf{e}_i) = \mathbf{v}_i$. Por tanto, la i-ésima columna de la matriz A que se busca está dada por

$$\begin{aligned}(S \circ T \circ R^{-1})(\mathbf{e}_i) &= S(T(R^{-1}(\mathbf{e}_i))) \\ &= S(T(\mathbf{v}_i)) \\ &= [T(\mathbf{v}_i)]_{\mathcal{C}}\end{aligned}$$

que es el vector coordenado de $T(\mathbf{v}_i)$ con respecto a la base $\mathcal{C}$ de W.

Este análisis se resume como teorema.

Teorema 6.26

Sean V y W dos espacios vectoriales con dimensión finita con bases $\mathcal{B}$ y $\mathcal{C}$, respectivamente, donde $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Si $T: V \to W$ es una transformación lineal, entonces la matriz A *de* $m \times n$ definida por

$$A = [[T(\mathbf{v}_1)]_{\mathcal{C}} \mid [T(\mathbf{v}_2)]_{\mathcal{C}} \mid \cdots \mid [T(\mathbf{v}_n)]_{\mathcal{C}}]$$

satisface

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

para todo vector $\mathbf{v}$ en V.

La matriz A en el Teorema 6.26 se llama ***matriz de T con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$.*** A continuación se ilustra la relación. (Recuerde que T_A denota multiplicación por A.)

$$\begin{array}{ccc} \mathbf{v} & \xrightarrow{T} & T(\mathbf{v}) \\ \downarrow & & \downarrow \\ [\mathbf{v}]_{\mathcal{B}} & \xrightarrow{T_A} & A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}} \end{array}$$

Comentarios

- La matriz de una transformación lineal T con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$ en ocasiones se denota mediante $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$. Note la dirección de la flecha: derecha a izquierda (no izquierda a derecha, como para $T: V \to W$). Con esta notación, la ecuación final en el Teorema 6.26 se convierte en

$$[T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

Observe que las $\mathcal{B}$ en los subíndices aparecen lado a lado y parecen "cancelarse" mutuamente. Con palabras, esta ecuación dice: "la matriz para T multiplicado por el vector coordenado para $\mathbf{v}$ produce el vector coordenado para $T(\mathbf{v})$."

En el caso especial donde $V = W$ y $\mathcal{B} = \mathcal{C}$, se escribe $[T]_{\mathcal{B}}$ (en lugar de $[T]_{\mathcal{B}\leftarrow\mathcal{B}}$). Entonces el Teorema 6.26 establece que

$$[T]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{B}}$$

- La matriz de una transformación lineal con respecto a las bases dadas es única. Esto es, para todo vector $\mathbf{v}$ en V, existe solamente *una* matriz A con la propiedad especificada por el Teorema 6.26, a saber,

$$A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$$

(En el ejercicio 39 se le pide probar esto.)

- El diagrama que sigue al Teorema 6.26 en ocasiones se conoce como *diagrama conmutativo* porque puede iniciar en la esquina superior izquierda con el vector **v** y llegar a $[T(\mathbf{v})]_{\mathcal{C}}$ en la esquina inferior derecha en dos formas diferentes, pero equivalentes. Si, como antes, los mapeos coordenados que mapean **v** a $[\mathbf{v}]_{\mathcal{B}}$ y **w** a $[\mathbf{w}]_{\mathcal{C}}$ se denotan R y S, respectivamente, entonces esta "conmutatividad" puede resumirse como

$$S \circ T = T_A \circ R$$

La razón para el término *conmutativo* se vuelve más clara cuando $V = W$ y $\mathcal{B} = \mathcal{C}$, porque entonces $R = S$ también, y se tiene

$$R \circ T = T_A \circ R$$

lo que sugiere que el mapeo coordenado R conmuta con la transformación lineal T (siempre que se use la versión matricial de T, a saber $T_A = T_{[T]_{\mathcal{B}}}$ donde se requiere).

- La matriz $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ depende del *orden* de los vectores en las bases $\mathcal{B}$ y $\mathcal{C}$. Al reordenar los vectores dentro de cualquiera de las bases afectará la matriz $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$. [Vea el ejemplo 6.77(b).]

Ejemplo 6.76

Sea $T : \mathbb{R}^3 \to \mathbb{R}^2$ la transformación lineal definida por

$$T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - 2y \\ x + y - 3z \end{bmatrix}$$

y sean $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ y $\mathcal{C} = \{\mathbf{e}_2, \mathbf{e}_1\}$ bases para $\mathbb{R}^3$ y $\mathbb{R}^2$, respectivamente. Encuentre la matriz de T con respecto a $\mathcal{B}$ y $\mathcal{C}$ y verifique el Teorema 6.26 para $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$.

Solución Primero calcule

$$T(\mathbf{e}_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad T(\mathbf{e}_2) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad T(\mathbf{e}_3) = \begin{bmatrix} 0 \\ -3 \end{bmatrix}$$

A continuación se necesitan sus vectores coordenados con respecto a $\mathcal{C}$. Dado que

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{e}_2 + \mathbf{e}_1, \quad \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \mathbf{e}_2 - 2\mathbf{e}_1, \quad \begin{bmatrix} 0 \\ -3 \end{bmatrix} = -3\mathbf{e}_2 + 0\mathbf{e}_1$$

se tiene

$$[T(\mathbf{e}_1)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad [T(\mathbf{e}_2)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad [T(\mathbf{e}_3)]_{\mathcal{C}} = \begin{bmatrix} -3 \\ 0 \end{bmatrix}$$

Por tanto, la matriz de T con respecto a $\mathcal{B}$ y $\mathcal{C}$ es

$$\begin{aligned} A = [T]_{\mathcal{C}\leftarrow\mathcal{B}} &= [[T(\mathbf{e}_1)]_{\mathcal{C}} \;\; [T(\mathbf{e}_2)]_{\mathcal{C}} \;\; [T(\mathbf{e}_3)]_{\mathcal{C}}] \\ &= \begin{bmatrix} 1 & 1 & -3 \\ 1 & -2 & 0 \end{bmatrix} \end{aligned}$$

Para verificar el Teorema 6.26 para **v**, calcule primero

$$T(\mathbf{v}) = T\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -5 \\ 10 \end{bmatrix}$$

Entonces

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix}$$

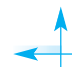
y

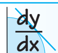
$$[T(\mathbf{v})]_{\mathcal{C}} = \begin{bmatrix} -5 \\ 10 \end{bmatrix}_{\mathcal{C}} = \begin{bmatrix} 10 \\ -5 \end{bmatrix}$$


(Compruebe esto.)

Al usar todos estos hechos, se confirma que

$$A[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & -3 \\ 1 & -2 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 10 \\ -5 \end{bmatrix} = [T(\mathbf{v})]_{\mathcal{C}}$$





Ejemplo 6.77

Sea $D : \mathscr{P}_3 \to \mathscr{P}_2$ el operador diferencial $D(p(x)) = p'(x)$. Sean $\mathcal{B} = \{1, x, x^2, x^3\}$ y $\mathcal{C} = \{1, x, x^2\}$ bases para $\mathscr{P}_3$ y $\mathscr{P}_2$, respectivamente.

(a) Encuentre la matriz A de D con respecto a $\mathcal{B}$ y $\mathcal{C}$.

(b) Encuentre la matriz A' de D con respecto a $\mathcal{B}'$ y $\mathcal{C}$, donde $\mathcal{B}' = \{x^3, x^2, x, 1\}$.

(c) Con el inciso (a), calcule $D(5 - x + 2x^3)$ y $D(a + bx + cx^2 + dx^3)$ para verificar el Teorema 6.26.

Solución Note primero que $D(a + bx + cx^2 + dx^3) = b + 2cx + 3dx^2$. (Vea el ejemplo 6.60.)

(a) Dado que las imágenes de la base $\mathcal{B}$ bajo D son $D(1) = 0$, $D(x) = 1$, $D(x^2) = 2x$ y $D(x^3) = 3x^2$, sus vectores coordenados con respeto a $\mathcal{C}$ son

$$[D(1)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad [D(x)]_{\mathcal{C}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [D(x^2)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix}, \quad [D(x^3)]_{\mathcal{C}} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}$$

En consecuencia,

$$\begin{aligned} A = [D]_{\mathcal{C}\leftarrow\mathcal{B}} &= [[D(1)]_{\mathcal{C}} \mid [D(x)]_{\mathcal{C}} \mid [D(x^2)]_{\mathcal{C}} \mid [D(x^3)]_{\mathcal{C}}] \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \end{aligned}$$

(b) Puesto que la base $\mathcal{B}'$ es justo $\mathcal{B}$ en el orden *inverso*, se ve que

$$\begin{aligned} A' = [D]_{\mathcal{C}\leftarrow\mathcal{B}'} &= [[D(x^3)]_{\mathcal{C}} \mid [D(x^2)]_{\mathcal{C}} \mid [D(x)]_{\mathcal{C}} \mid [D(1)]_{\mathcal{C}}] \\ &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

(Esto demuestra que el *orden* de los vectores en las bases $\mathcal{B}$ y $\mathcal{C}$ afecta la matriz de una transformación con respecto a dichas bases.)

(c) Primero calcule directamente $D(5 - x + 2x^3) = -1 + 6x^2$ y obtenga el vector coordenado

$$[D(5 - x + 2x^3)]_{\mathcal{C}} = [-1 + 6x^2]_{\mathcal{C}} = \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix}$$

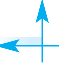
Por otra parte,

$$[5 - x + 2x^3]_{\mathcal{B}} = \begin{bmatrix} 5 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$

de modo que

$$A[5 - x + 2x^3]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ -1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix} = [D(5 - x + 2x^3)]_{\mathcal{C}}$$

lo que concuerda con el Teorema 6.26. La demostración del caso general se deja como ejercicio.



Dado que la transformación lineal en el ejemplo 6.77 es fácil de usar directamente, en realidad no hay ventaja al usar la matriz de esta transformación para hacer los cálculos. Sin embargo, en otros ejemplos, especialmente los laboriosos, el esquema matricial puede ser más simple, pues es muy adecuado para su implementación en computadora. El ejemplo 6.78 ilustra la idea básica detrás de este enfoque indirecto.

Ejemplo 6.78

Sea $T : \mathscr{P}_2 \to \mathscr{P}_2$ la transformación lineal definida por

$$T(p(x)) = p(2x - 1)$$

(a) Encuentre la matriz de T con respecto a $\mathcal{E} = \{1, x, x^2\}$.

(b) Calcule $T(3 + 2x - x^2)$ indirectamente, use el inciso (a).

Solución (a) Se ve que

$$T(1) = 1, \quad T(x) = 2x - 1, \quad T(x^2) = (2x - 1)^2 = 1 - 4x + 4x^2$$

de modo que los vectores coordenados son

$$[T(1)]_{\mathcal{E}} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad [T(x)]_{\mathcal{E}} = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \quad [T(x^2)]_{\mathcal{E}} = \begin{bmatrix} 1 \\ -4 \\ 4 \end{bmatrix}$$

En consecuencia,

$$[T]_{\mathcal{E}} = [[T(1)]_{\mathcal{E}} \mid [T(x)]_{\mathcal{E}} \mid [T(x^2)]_{\mathcal{E}}] = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix}$$

(b) El Teorema 6.26 se aplica del modo siguiente: el vector coordenado de $p(x) = 3 + 2x - x^2$ con respecto a $\mathcal{E}$ es

$$[p(x)]_{\mathcal{E}} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$$

Por tanto, por el Teorema 6.26,

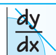
$$\begin{aligned} [T(3 + 2x - x^2)]_{\mathcal{E}} &= [T(p(x))]_{\mathcal{E}} \\ &= [T]_{\mathcal{E}}[p(x)]_{\mathcal{E}} \\ &= \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ -4 \end{bmatrix} \end{aligned}$$



Se concluye que $T(3 + 2x - x^2) = 0 \cdot 1 + 8 \cdot x - 4 \cdot x^2 = 8x - 4x^2$. [Verifique esto al calcular directamente $T(3 + 2x - x^2) = 3 + 2(2x - 1) - (2x - 1)^2$.]



La matriz de una transformación lineal a veces puede usarse en formas sorprendentes. El ejemplo 6.79 muestra su aplicación a un problema de cálculo tradicional.



Ejemplo 6.79



Sea $\mathcal{D}$ el espacio vectorial de todas las funciones derivables. Considere el subespacio W de $\mathcal{D}$ dado por $W = \text{gen}(e^{3x}, xe^{3x}, x^2e^{3x})$. Puesto que el conjunto $\mathcal{B} = \{e^{3x}, xe^{3x}, x^2e^{3x}\}$ es linealmente independiente (¿por qué?), es una base para W.

(a) Demuestre que el operador diferencial D mapea a W en sí mismo.

(b) Encuentre la matriz de D con respecto a $\mathcal{B}$.

(c) Calcule indirectamente la derivada de $5e^{3x} + 2xe^{3x} - x^2e^{3x}$ con el Teorema 6.26 y verifíquela usando el inciso (a).

Solución (a) Al aplicar D a un elemento general de W, se ve que

$$D(ae^{3x} + bxe^{3x} + cx^2e^{3x}) = (3a + b)e^{3x} + (3b + 2c)xe^{3x} + 3cx^2e^{3x}$$



(compruebe esto), que de nuevo está en W.

(b) Al usar la fórmula del inciso (a), se ve que

$$D(e^{3x}) = 3e^{3x}, \quad D(xe^{3x}) = e^{3x} + 3xe^{3x}, \quad D(x^2e^{3x}) = 2xe^{3x} + 3x^2e^{3x}$$

de modo que

$$[D(e^{3x})]_{\mathcal{B}} = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}, \quad [D(xe^{3x})]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}, \quad [D(x^2e^{3x})]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}$$

Se tiene que

$$[D]_{\mathcal{B}} = [[D(e^{3x})]_{\mathcal{B}} \mid [D(xe^{3x})]_{\mathcal{B}} \mid [D(x^2e^{3x})]_{\mathcal{B}}] = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

(c) Para $f(x) = 5e^{3x} + 2xe^{3x} - x^2e^{3x}$, se ve por inspección que

$$[f(x)]_{\mathcal{B}} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

En consecuencia, por el Teorema 6.26, se tiene

$$[D(f(x))]_{\mathcal{B}} = [D]_{\mathcal{B}}[f(x)]_{\mathcal{B}} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 17 \\ 4 \\ -3 \end{bmatrix}$$

lo que, a su vez, implica que $f'(x) = D(f(x)) = 17e^{3x} + 4xe^{3x} - 3x^2e^{3x}$, en concordancia con la fórmula del inciso (a).

Comentario El punto del ejemplo 6.79 no es que este método sea más sencillo que la derivación directa. De hecho, una vez establecida la fórmula en el inciso (a), hay poco que hacer. Lo que es significativo es que los métodos matriciales pueden usarse en lo que, en la superficie, parece ser un problema de cálculo. Esta idea se explorará aún más en el ejemplo 6.83.

Ejemplo 6.80

Sea V un espacio vectorial n-dimensional y sea I la transformación identidad en V. ¿Cuál es la matriz de I con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$ de V, si $\mathcal{B} = \mathcal{C}$ (incluido el orden de los vectores base)? ¿Qué pasa si $\mathcal{B} \neq \mathcal{C}$?

Solución Sea $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. Entonces $I(\mathbf{v}_1) = \mathbf{v}_1, \ldots, I(\mathbf{v}_n) = \mathbf{v}_n$, de modo que

$$[I(\mathbf{v}_1)]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{e}_1, \quad [I(\mathbf{v}_2)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{e}_2, \quad \ldots, \quad [I(\mathbf{v}_n)]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} = \mathbf{e}_n$$

y, si $\mathcal{B} = \mathcal{C}$,

$$\begin{aligned} [I]_{\mathcal{B}} &= [[I(\mathbf{v}_1)]_{\mathcal{B}} \mid [I(\mathbf{v}_2)]_{\mathcal{B}} \mid \cdots \mid [I(\mathbf{v}_n)]_{\mathcal{B}}] \\ &= [\mathbf{e}_1 \mid \mathbf{e}_2 \mid \cdots \mid \mathbf{e}_n] \\ &= I_n \end{aligned}$$

la matriz identidad de $n \times n$. (Esto es lo que esperaba, ¿o no?)

En el caso $\mathcal{B} \neq \mathcal{C}$, se tiene

$$[I(\mathbf{v}_1)]_{\mathcal{C}} = [\mathbf{v}_1]_{\mathcal{C}}, \quad \ldots, \quad [I(\mathbf{v}_n)]_{\mathcal{C}} = [\mathbf{v}_n]_{\mathcal{C}}$$

de modo que

$$\begin{aligned} [I]_{\mathcal{C}\leftarrow\mathcal{B}} &= [[\mathbf{v}_1]_{\mathcal{C}} \mid \cdots \mid [\mathbf{v}_n]_{\mathcal{C}}] \\ &= P_{\mathcal{C}\leftarrow\mathcal{B}} \end{aligned}$$

la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$.

Matrices de transformaciones lineales compuesta e inversa

Ahora se generalizarán los Teoremas 3.32 y 3.33 para obtener un teorema que permitirá encontrar fácilmente la inversa de una transformación lineal entre espacios vectoriales de dimensión finita (si existe).

Teorema 6.27 Sean U, V y W espacios vectoriales de dimensión finita con bases $\mathcal{B}$, $\mathcal{C}$ y $\mathcal{D}$, respectivamente. Sean $T: U \to V$ y $S: V \to W$ transformaciones lineales. Entonces

$$[S \circ T]_{\mathcal{D}\leftarrow\mathcal{B}} = [S]_{\mathcal{D}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}}$$

Comentarios

- En palabras, este teorema dice: "la matriz de la compuesta es el producto de las matrices".
- Note cómo los "subíndices interiores" $\mathcal{C}$ deben coincidir y parecen cancelarse mutuamente, lo que deja los "subíndices exteriores" en la forma $\mathcal{D} \leftarrow \mathcal{B}$.

Demostración Se demostrará que las columnas correspondientes de las matrices $[S \circ T]_{\mathcal{D}\leftarrow\mathcal{B}}$ y $[S]_{\mathcal{D}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ son iguales. Sea $\mathbf{v}_i$ el i-ésimo vector base en $\mathcal{B}$. Entonces la i-ésima columna de $[S \circ T]_{\mathcal{D}\leftarrow\mathcal{B}}$ es

$$\begin{aligned} [(S \circ T)(\mathbf{v}_i)]_{\mathcal{D}} &= [S(T(\mathbf{v}_i))]_{\mathcal{D}} \\ &= [S]_{\mathcal{D}\leftarrow\mathcal{C}}[T(\mathbf{v}_i)]_{\mathcal{C}} \\ &= [S]_{\mathcal{D}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}_i]_{\mathcal{B}} \end{aligned}$$

por dos aplicaciones del Teorema 6.26. Pero $[\mathbf{v}_i]_{\mathcal{B}} = \mathbf{e}_i$ (¿por qué?), de modo que

$$[S]_{\mathcal{D}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}}[\mathbf{v}_i]_{\mathcal{B}} = [S]_{\mathcal{D}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}}\mathbf{e}_i$$

es la i-ésima columna de la matriz $[S]_{\mathcal{D}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}}$. Por tanto, las i-ésimas columnas de $[S \circ T]_{\mathcal{D}\leftarrow\mathcal{B}}$ y $[S]_{\mathcal{D}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ son iguales, como se quería demostrar.

Ejemplo 6.81 Use métodos matriciales para calcular $(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix}$ para las transformaciones lineales S y T del ejemplo 6.56.

Solución Recuerde que $T: \mathbb{R}^2 \to \mathscr{P}_1$ y $S: \mathscr{P}_1 \to \mathscr{P}_2$ están definidas por

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x \quad \text{y} \quad S(a + bx) = ax + bx^2$$

Al elegir las bases estándar $\mathcal{E}$, $\mathcal{E}'$ y $\mathcal{E}''$ para $\mathbb{R}^2$, $\mathscr{P}_1$ y $\mathscr{P}_2$, respectivamente, se ve que

$$[T]_{\mathcal{E}'\leftarrow\mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad [S]_{\mathcal{E}''\leftarrow\mathcal{E}'} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(Verifique esto.) Por el Teorema 6.27, la matriz de $S \circ T$ con respecto a $\mathcal{E}$ y $\mathcal{E}''$ es

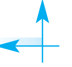
$$\begin{aligned} [(S \circ T)]_{\mathcal{E}''\leftarrow\mathcal{E}} &= [S]_{\mathcal{E}''\leftarrow\mathcal{E}'}[T]_{\mathcal{E}'\leftarrow\mathcal{E}} \\ &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \end{aligned}$$

Por tanto, por el Teorema 6.26,

$$\left[(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix}\right]_{\mathcal{E}''} = [(S \circ T)]_{\mathcal{E}'' \leftarrow \mathcal{E}} \begin{bmatrix} a \\ b \end{bmatrix}_{\mathcal{E}}$$

$$= \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ a \\ a + b \end{bmatrix}$$

En consecuencia, $(S \circ T)\begin{bmatrix} a \\ b \end{bmatrix} = ax + (a + b)x^2$, que concuerda con la solución al ejemplo 6.56.



En el Teorema 6.24 se demostró que una transformación lineal es invertible si y sólo si es inyectiva y sobreyectiva (esto es, si es un isomorfismo). Cuando los espacios vectoriales involucrados son de dimensión finita, puede usar los métodos matriciales desarrollados para encontrar la inversa de tal transformación lineal.

Teorema 6.28

Sea $T : V \to W$ una transformación lineal entre espacios vectoriales V y W de n-dimensiones, y sean $\mathcal{B}$ y $\mathcal{C}$ bases para V y W, respectivamente. Entonces T es invertible si y sólo si la matriz $[T]_{\mathcal{C} \leftarrow \mathcal{B}}$ es invertible. En este caso,

$$([T]_{\mathcal{C} \leftarrow \mathcal{B}})^{-1} = [T^{-1}]_{\mathcal{B} \leftarrow \mathcal{C}}$$

Demostración Observe que las matrices de T y T^{-1} (si existen) son de $n \times n$. Si T es invertible, entonces $T^{-1} \circ T = I_V$. Al aplicar el Teorema 6.27, se tiene

$$I_n = [I_V]_{\mathcal{B}} = [T^{-1} \circ T]_{\mathcal{B}}$$
$$= [T^{-1}]_{\mathcal{B} \leftarrow \mathcal{C}}[T]_{\mathcal{C} \leftarrow \mathcal{B}}$$

Esto demuestra que $[T]_{\mathcal{C} \leftarrow \mathcal{B}}$ es invertible y que $([T]_{\mathcal{C} \leftarrow \mathcal{B}})^{-1} = [T^{-1}]_{\mathcal{B} \leftarrow \mathcal{C}}$.

Por el contrario, suponga que $A = [T]_{\mathcal{C} \leftarrow \mathcal{B}}$ es invertible. Para demostrar que T es invertible, es suficiente demostrar que $\ker(T) = \{\mathbf{0}\}$. (¿Por qué?) Para este fin, sea $\mathbf{v}$ que está en el kernel de T. Entonces $T(\mathbf{v}) = \mathbf{0}$, de modo que

$$A[\mathbf{v}]_{\mathcal{B}} = [T]_{\mathcal{C} \leftarrow \mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}} = [\mathbf{0}]_{\mathcal{C}} = \mathbf{0}$$

lo que significa que $[\mathbf{v}]_{\mathcal{B}}$ está en el espacio nulo de la matriz invertible A. Por el Teorema fundamental, esto implica que $[\mathbf{v}]_{\mathcal{B}} = \mathbf{0}$, lo cual, a su vez, implica que $\mathbf{v} = \mathbf{0}$, como se requiere.

Ejemplo 6.82

En el ejemplo 6.70 se demostró que la transformación lineal $T : \mathbb{R}^2 \to \mathscr{P}_1$ definida por

$$T\begin{bmatrix} a \\ b \end{bmatrix} = a + (a + b)x$$

era inyectiva y sobreyectiva, y por tanto invertible. Encuentre T^{-1}.

Solución En el ejemplo 6.81 se encontró que la matriz de T con respecto a las bases estándar $\mathcal{E}$ y $\mathcal{E}'$ para $\mathbb{R}^2$ y $\mathscr{P}_1$, respectivamente, era

$$[T]_{\mathcal{E}'\leftarrow\mathcal{E}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

Por el Teorema 6.28 se tiene que la matriz de T^{-1} con respecto a $\mathcal{E}'$ y $\mathcal{E}$ es

$$[T^{-1}]_{\mathcal{E}\leftarrow\mathcal{E}'} = ([T]_{\mathcal{E}'\leftarrow\mathcal{E}})^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$$

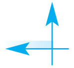
Por el Teorema 6.26,

$$\begin{aligned} [T^{-1}(a + bx)]_{\mathcal{E}} &= [T^{-1}]_{\mathcal{E}\leftarrow\mathcal{E}'}[a + bx]_{\mathcal{E}'} \\ &= \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} \\ &= \begin{bmatrix} a \\ b - a \end{bmatrix} \end{aligned}$$

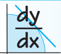
Esto significa que

$$T^{-1}(a + bx) = a\mathbf{e}_1 + (b - a)\mathbf{e}_2 = \begin{bmatrix} a \\ b - a \end{bmatrix}$$

(Note que la elección de la base estándar hace este cálculo virtualmente irrelevante.)



El siguiente ejemplo, una continuación del ejemplo 6.79, muestra que las matrices pueden usarse en ciertos problemas de integración en cálculo. La integral específica que se considerará por lo general se evalúa en cursos de cálculo mediante dos aplicaciones de la integración por partes. Contraste este procedimiento con el método propuesto.



Ejemplo 6.83

Demuestre que el operador diferencial, restringido al subespacio $W = \text{gen}(e^{3x}, xe^{3x}, x^2e^{3x})$ de $\mathscr{D}$, es invertible, y use este hecho para encontrar la integral

$$\int x^2e^{3x}\,dx$$

Solución En el ejemplo 6.79 se encontró que la matriz de D con respecto a la base $\mathcal{B} = \{e^{3x}, xe^{3x}, x^2e^{3x}\}$ de W era

$$[D]_{\mathcal{B}} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}$$

Por tanto, por el Teorema 6.28, D es invertible sobre W, y la matriz de D^{-1} es

$$[D^{-1}]_{\mathcal{B}} = ([D]_{\mathcal{B}})^{-1} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 3 & 2 \\ 0 & 0 & 3 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{3} & -\frac{1}{9} & \frac{2}{27} \\ 0 & \frac{1}{3} & -\frac{2}{9} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}$$

Puesto que la integración es *antiderivación*, esta es la matriz correspondiente a la integración sobre W. Se desea integrar la función x^2e^{3x} cuyo vector coordenado es

$$[x^2e^{3x}]_{\mathcal{B}} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

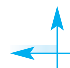
En consecuencia, por el Teorema 6.26,

$$\begin{aligned} \left[\int x^2e^{3x}\,dx\right]_{\mathcal{B}} &= [D^{-1}(x^2e^{3x})]_{\mathcal{B}} \\ &= [D^{-1}]_{\mathcal{B}}[x^2e^{3x}]_{\mathcal{B}} \\ &= \begin{bmatrix} \frac{1}{3} & -\frac{1}{9} & \frac{2}{27} \\ 0 & \frac{1}{3} & -\frac{2}{9} \\ 0 & 0 & \frac{1}{3} \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{2}{27} \\ -\frac{2}{9} \\ \frac{1}{3} \end{bmatrix} \end{aligned}$$

Se concluye que

$$\int x^2e^{3x}\,dx = \tfrac{2}{27}e^{3x} - \tfrac{2}{9}xe^{3x} + \tfrac{1}{3}x^2e^{3x}$$

(Para estar plenamente correctos, es necesario sumar una constante de integración. Aquí no se muestra porque se trabaja con transformaciones *lineales*, que deben enviar vectores cero a vectores cero, lo que fuerza que la constante de integración también sea cero.)



Advertencia En general, la derivación *no* es una transformación invertible. (Vea el ejercicio 22.) Lo que demuestra el ejemplo precedente es que, restringida de manera adecuada, en ocasiones sí lo es. Los ejercicios 27-30 exploran más esta idea.

Cambio de base y similitud

Suponga que $T: V \to V$ es una transformación lineal y $\mathcal{B}$ y $\mathcal{C}$ son dos bases diferente para V. Es natural preguntar cómo se relacionan las matrices $[T]_{\mathcal{B}}$ y $[T]_{\mathcal{C}}$, si es que lo hacen. Es evidente que la respuesta a esta pregunta es bastante satisfactoria y se relaciona con algunas preguntas que se consideraron por primera vez en el capítulo 4.

La figura 6.12 sugiere una forma de abordar este problema. Al seguir las flechas alrededor del diagrama, desde la esquina superior izquierda hasta la esquina inferior derecha, en dos caminos diferentes, pero equivalentes, se demuestra que $I \circ T = T \circ I$, algo que ya se sabía, pues ambas son iguales a T. Sin embargo, si la versión "superior" de T es

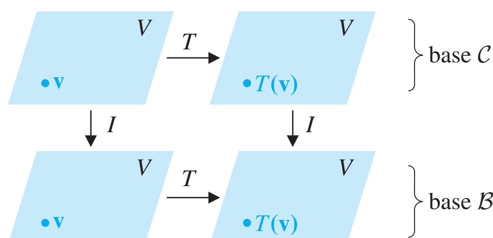


Figura 6.12
$I \circ T = T \circ I$

con respecto a la base $\mathcal{C}$ y la versión "inferior" es con respecto a $\mathcal{B}$, entonces $T = I \circ T = T \circ I$ es con respecto a $\mathcal{C}$ en su dominio y con respecto a $\mathcal{B}$ en su codominio. Por tanto, la matriz de T en este caso es $[T]_{\mathcal{B}\leftarrow\mathcal{C}}$. Pero

$$[T]_{\mathcal{B}\leftarrow\mathcal{C}} = [I \circ T]_{\mathcal{B}\leftarrow\mathcal{C}} = [I]_{\mathcal{B}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{C}}$$

y

$$[T]_{\mathcal{B}\leftarrow\mathcal{C}} = [T \circ I]_{\mathcal{B}\leftarrow\mathcal{C}} = [T]_{\mathcal{B}\leftarrow\mathcal{B}}[I]_{\mathcal{B}\leftarrow\mathcal{C}}$$

Así que, $[I]_{\mathcal{B}\leftarrow\mathcal{C}}[T]_{\mathcal{C}\leftarrow\mathcal{C}} = [T]_{\mathcal{B}\leftarrow\mathcal{B}}[I]_{\mathcal{B}\leftarrow\mathcal{C}}$.

A partir del ejemplo 6.80, se sabe que $[I]_{\mathcal{B}\leftarrow\mathcal{C}} = P_{\mathcal{B}\leftarrow\mathcal{C}}$, la matriz (invertible) de cambio de base de $\mathcal{C}$ a $\mathcal{B}$. Si esta matriz se denota P, entonces también se tiene

$$P^{-1} = (P_{\mathcal{B}\leftarrow\mathcal{C}})^{-1} = P_{\mathcal{C}\leftarrow\mathcal{B}}$$

Con esta notación,

$$P[T]_{\mathcal{C}\leftarrow\mathcal{C}} = [T]_{\mathcal{B}\leftarrow\mathcal{B}}P$$

de modo que $\quad [T]_{\mathcal{C}\leftarrow\mathcal{C}} = P^{-1}[T]_{\mathcal{B}\leftarrow\mathcal{B}}P \quad$ o $\quad [T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P$

Por tanto, las matrices $[T]_{\mathcal{B}}$ y $[T]_{\mathcal{C}}$ son similares, en la terminología de la sección 4.4.

La discusión anterior se resume como teorema.

Teorema 6.29

Sea V un espacio vectorial de dimensión finita con bases $\mathcal{B}$ y $\mathcal{C}$ y sea $T: V \to V$ una transformación lineal. Entonces

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}P$$

donde P es la matriz de cambio de base de $\mathcal{C}$ a $\mathcal{B}$.

Comentario Como auxiliar para recordar que P debe ser la matriz de cambio de base de $\mathcal{C}$ a $\mathcal{B}$ y no de $\mathcal{B}$ a $\mathcal{C}$, es instructivo observar lo que dice el Teorema 6.29 cuando se escribe con todo detalle. Como se muestra a continuación, los "subíndices interiores" deben ser iguales (todos $\mathcal{B}$) y deben parecer cancelarse, lo que deja los "subíndices exteriores", que son ambos $\mathcal{C}$.

$$[T]_{C\leftarrow C} = P_{C\leftarrow B}[T]_{B\leftarrow B}P_{B\leftarrow C}$$

Igual Igual

Igual

Con frecuencia se usa el Teorema 6.29 cuando se intenta encontrar una base con respecto a la cual la matriz de una transformación lineal es particularmente simple. Por ejemplo, es posible preguntar si existe una base $\mathcal{C}$ de V tal que la matriz $[T]_{\mathcal{C}}$ de $T: V \to V$ es una matriz diagonal. El ejemplo 6.84 ilustra esta aplicación.

Ejemplo 6.84

Sea $T: \mathbb{R}^2 \to \mathbb{R}^2$ definida mediante

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 3y \\ 2x + 2y \end{bmatrix}$$

Si es posible, encuentre una base $\mathcal{C}$ para $\mathbb{R}^2$ tal que la matriz de T con respecto a $\mathcal{C}$ sea diagonal.

Solución La matriz de T con respecto a la base estándar $\mathcal{E}$ es

$$[T]_{\mathcal{E}} = \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}$$

Esta matriz es diagonalizable, como se vio en el ejemplo 4.24. De hecho, si

$$P = \begin{bmatrix} 1 & 3 \\ 1 & -2 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 4 & 0 \\ 0 & -1 \end{bmatrix}$$

entonces $P^{-1}[T]_{\mathcal{E}}P = D$. Si se hace $\mathcal{C}$ la base de $\mathbb{R}^2$ consistente de las columnas de P, entonces P es la matriz de cambio de base $P_{\mathcal{E}\leftarrow\mathcal{C}}$ de $\mathcal{C}$ a $\mathcal{E}$. Por el Teorema 6.29,

$$[T]_{\mathcal{C}} = P^{-1}[T]_{\mathcal{E}}P = D$$

de modo que la matriz de T con respecto a la base $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \end{bmatrix} \right\}$ es diagonal.

Comentarios

- Es fácil comprobar que la solución anterior es correcta al calcular directamente $[T]_{\mathcal{C}}$. Se encuentra que

$$T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0\begin{bmatrix} 3 \\ -2 \end{bmatrix} \quad \text{y} \quad T\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Por tanto, los vectores coordenados que forman las columnas de $[T]_{\mathcal{C}}$ son

$$\left[T\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right]_{\mathcal{C}} = \begin{bmatrix} 4 \\ 0 \end{bmatrix} \quad \text{y} \quad \left[T\begin{bmatrix} 3 \\ -2 \end{bmatrix}\right]_{\mathcal{C}} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

en concordancia con la solución anterior.

- El procedimiento general para un problema como el ejemplo 6.84 es tomar la matriz estándar $[T]_{\mathcal{E}}$ y determinar si es diagonalizable al encontrar bases para sus eigenespacios, como en el capítulo 4. Entonces la solución procede exactamente como en el ejemplo anterior.

El ejemplo 6.84 motiva la siguiente definición.

Definición Sea V un espacio vectorial de dimensión finita y sea $T: V \to V$ una transformación lineal. Entonces T se llama ***diagonalizable*** si existe una base $\mathcal{C}$ para V tal que la matriz $[T]_{\mathcal{C}}$ sea una matriz diagonal.

No es difícil demostrar que, si $\mathcal{B}$ es *cualquier* base para V, entonces T es diagonalizable si y sólo si la matriz $[T]_{\mathcal{B}}$ es diagonalizable. En esencia esto es lo que se hizo, para un caso especial, en el último ejemplo. En el ejercicio 42 se le pide probar este resultado en general.

En ocasiones es más sencillo escribir la matriz de una transformación lineal con respecto a una base "no estándar". Entonces puede invertir el proceso del ejemplo 6.84 para encontrar la matriz estándar. Esta idea se ilustra al revisar nuevamente el ejemplo 3.59.

Ejemplo 6.85

Sea ℓ la recta que pasa por el origen en $\mathbb{R}^2$ con vector de director $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$. Encuentre la matriz estándar de la proyección sobre ℓ.

Solución Sea T la proyección. No hace daño suponer que $\mathbf{d}$ es un vector unitario (es decir, $d_1^2 + d_2^2 = 1$), pues cualquier múltiplo de $\mathbf{d}$ distinto de cero puede servir como un vector director para ℓ. Sea $\mathbf{d}' = \begin{bmatrix} -d_2 \\ d_1 \end{bmatrix}$ de modo que $\mathbf{d}$ y $\mathbf{d}'$ son ortogonales. Dado que $\mathbf{d}'$ también es un vector unitario, el conjunto $\mathcal{D} = \{\mathbf{d}, \mathbf{d}'\}$ es una base ortonormal para $\mathbb{R}^2$.

Como muestra la figura 6.13, $T(\mathbf{d}) = \mathbf{d}$ y $T(\mathbf{d}') = \mathbf{0}$. Por tanto,

$$[T(\mathbf{d})]_{\mathcal{D}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad [T(\mathbf{d}')]_{\mathcal{D}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

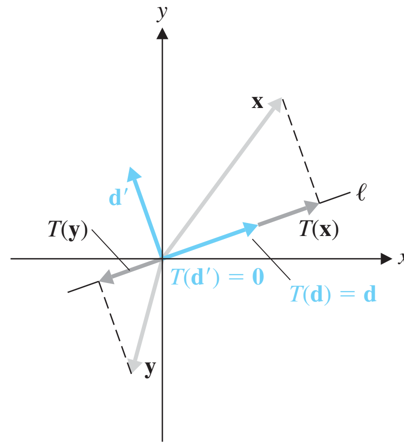


Figura 6.13
Proyección sobre ℓ

así
$$[T]_{\mathcal{D}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

La matriz de cambio de base de $\mathcal{D}$ a la base estándar $\mathcal{E}$ es

$$P_{\mathcal{E}\leftarrow\mathcal{D}} = \begin{bmatrix} d_1 & -d_2 \\ d_2 & d_1 \end{bmatrix}$$

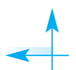
de modo que la matriz de cambio de base de $\mathcal{E}$ a $\mathcal{D}$ es

$$P_{\mathcal{D}\leftarrow\mathcal{E}} = (P_{\mathcal{E}\leftarrow\mathcal{D}})^{-1} = \begin{bmatrix} d_1 & -d_2 \\ d_2 & d_1 \end{bmatrix}^{-1} = \begin{bmatrix} d_1 & d_2 \\ -d_2 & d_1 \end{bmatrix}$$

Entonces, por el Teorema 6.29, la matriz estándar de T es

$$\begin{aligned} [T]_{\mathcal{E}} &= P_{\mathcal{E}\leftarrow\mathcal{D}}[T]_{\mathcal{D}}P_{\mathcal{D}\leftarrow\mathcal{E}} \\ &= \begin{bmatrix} d_1 & -d_2 \\ d_2 & d_1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} d_1 & d_2 \\ -d_2 & d_1 \end{bmatrix} \\ &= \begin{bmatrix} d_1^2 & d_1d_2 \\ d_1d_2 & d_2^2 \end{bmatrix} \end{aligned}$$

lo que concuerda con el inciso (b) del ejemplo 3.59.



Ejemplo 6.86

Sea $T : \mathscr{P}_2 \to \mathscr{P}_2$ la transformación lineal definida por

$$T(p(x)) = p(2x - 1)$$

(a) Encuentre la matriz de T con respecto a la base $\mathcal{B} = \{1 + x, 1 - x, x^2\}$ de $\mathscr{P}_2$.

(b) Demuestre que T es diagonalizable y encuentre una base $\mathcal{C}$ para $\mathscr{P}_2$ tal que $[T]_{\mathcal{C}}$ sea una matriz diagonalizable.

Solución (a) En el ejemplo 6.78 se encontró que la matriz de T con respecto a la base estándar $\mathcal{E} = \{1, x, x^2\}$ es

$$[T]_{\mathcal{E}} = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix}$$

La matriz de cambio de base de $\mathcal{B}$ a $\mathcal{E}$ es

$$P = P_{\mathcal{E} \leftarrow \mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Se tiene que la matriz de T con respecto a $\mathcal{B}$ es

$$\begin{aligned} [T]_{\mathcal{B}} &= P^{-1}[T]_{\mathcal{E}}P \\ &= \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 0 & 2 & -4 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & -\frac{3}{2} \\ -1 & 2 & \frac{5}{2} \\ 0 & 0 & 4 \end{bmatrix} \end{aligned}$$


(Compruébelo.)


(b) Los eigenvalores de $[T]_{\mathcal{E}}$ son 1, 2 y 4 (¿por qué?), así que, del Teorema 4.25, se sabe que $[T]_{\mathcal{E}}$ es diagonalizable. Los eigenvectores correspondientes a estos eigenvalores son

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

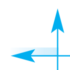
respectivamente. Por tanto, al hacer

$$P = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{y} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

se tiene $P^{-1}[T]_{\mathcal{E}}P = D$. Más aún, P es la matriz de cambio de base de una base $\mathcal{C}$ a $\mathcal{E}$ y las columnas de P son por tanto los vectores coordenados de $\mathcal{C}$ en términos de $\mathcal{E}$. Se concluye que

$$\mathcal{C} = \{1, -1 + x, 1 - 2x + x^2\}$$

y $[T]_{\mathcal{C}} = D$.



Las ideas anteriores pueden generalizarse para relacionar las matrices $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ y $[T]_{\mathcal{C}'\leftarrow\mathcal{B}'}$ de una transformación lineal $T: V \to W$, donde $\mathcal{B}$ y $\mathcal{B}'$ son bases para V y $\mathcal{C}$ y $\mathcal{C}'$ es base para W. (Vea el ejercicio 44.)

Esta sección concluye con una revisión del teorema fundamental de las matrices invertibles e incorpora algunos resultados de este capítulo.

Teorema 6.30 **El teorema fundamental de las matrices invertibles: versión 4**

Sea A una matriz de $n \times n$ y sea $T: V \to W$ una transformación lineal cuya matriz $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$ de V y W, respectivamente, es A. Las siguientes afirmaciones son equivalentes:

a. A es invertible.
b. $A\mathbf{x} = \mathbf{b}$ tiene una solución única para todo $\mathbf{b}$ en $\mathbb{R}^n$.
c. $A\mathbf{x} = \mathbf{0}$ tiene sólo la solución trivial.
d. La forma escalonada reducida por renglones de A es I_n.
e. A es un producto de matrices elementales.
f. $\text{rank}(A) = n$
g. $\text{nulidad}(A) = 0$
h. Los vectores columna de A son linealmente independientes.
i. Los vectores columna de A generan $\mathbb{R}^n$.
j. Los vectores columna de A forman una base para $\mathbb{R}^n$.
k. Los vectores renglón de A son linealmente independientes.
l. Los vectores renglón de A generan $\mathbb{R}^n$.
m. Los vectores renglón de A forman una base para $\mathbb{R}^n$.
n. $\det A \neq 0$
o. 0 no es un eigenvalor de A.
p. T es invertible.
q. T es inyectiva.
r. T es sobreyectiva.
s. $\ker(T) = \{\mathbf{0}\}$
t. $\text{rango}(T) = W$

Demostración La equivalencia (q)⇔(s) es el Teorema 6.20 y (r)⇔(t) es la definición de sobreyectiva. Dado que A es de $n \times n$, debe tener $\dim V = \dim W = n$. A partir de los Teoremas 6.21 y 6.24, se obtiene (p)⇔(q)⇔(r). Finalmente, los últimos cinco enunciados se conectan con los demás mediante el Teorema 6.28, que implica que (a)⇔(p).

Ejercicios 6.6

En los ejercicios 1-12, encuentre la matriz $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ de la transformación lineal $T: V \to W$ con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$ de V y W, respectivamente. Verifique el Teorema 6.26 para el vector $\mathbf{v}$ al calcular directamente $T(\mathbf{v})$ y usar el teorema.

1. $T: \mathcal{P}_1 \to \mathcal{P}_1$ definida mediante $T(a + bx) = b - ax$, $\mathcal{B} = \mathcal{C} = \{1, x\}$, $\mathbf{v} = p(x) = 4 + 2x$

2. $T: \mathcal{P}_1 \to \mathcal{P}_1$ definida mediante $T(a + bx) = b - ax$, $\mathcal{B} = \{1 + x,\ 1 - x\}$, $\mathcal{C} = \{1, x\}$, $\mathbf{v} = p(x) = 4 + 2x$

3. $T: \mathcal{P}_2 \to \mathcal{P}_2$ definida mediante $T(p(x)) = p(x + 2)$, $\mathcal{B} = \{1, x, x^2\}$, $\mathcal{C} = \{1, x + 2, (x + 2)^2\}$, $\mathbf{v} = p(x) = a + bx + cx^2$

4. $T: \mathscr{P}_2 \to \mathscr{P}_2$ definida mediante $T(p(x)) = p(x+2)$, $\mathcal{B} = \{1, x+2, (x+2)^2\}$, $\mathcal{C} = \{1, x, x^2\}$, $\mathbf{v} = p(x) = a + bx + cx^2$

5. $T: \mathscr{P}_2 \to \mathbb{R}^2$ definida mediante $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$, $\mathcal{B} = \{1, x, x^2\}$, $\mathcal{C} = \{\mathbf{e}_1, \mathbf{e}_2\}$, $\mathbf{v} = p(x) = a + bx + cx^2$

6. $T: \mathscr{P}_2 \to \mathbb{R}^2$ definida mediante $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$, $\mathcal{B} = \{x^2, x, 1\}$, $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$, $\mathbf{v} = p(x) = a + bx + cx^2$

7. $T: \mathbb{R}^2 \to \mathbb{R}^3$ definida mediante

$$T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a+2b \\ -a \\ b \end{bmatrix}, \quad \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ -1 \end{bmatrix} \right\},$$

$$\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}, \quad \mathbf{v} = \begin{bmatrix} -7 \\ 7 \end{bmatrix}$$

8. Repita el ejercicio 7 con $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$.

9. $T: M_{22} \to M_{22}$ definida mediante $T(A) = A^T$, $\mathcal{B} = \mathcal{C} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, $\mathbf{v} = A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

10. Repita el ejercicio 9 con $\mathcal{B} = \{E_{22}, E_{21}, E_{12}, E_{11}\}$ y $\mathcal{C} = \{E_{12}, E_{21}, E_{22}, E_{11}\}$.

11. $T: M_{22} \to M_{22}$ definida mediante $T(A) = AB - BA$, donde

$B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$, $\mathcal{B} = \mathcal{C} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$,

$\mathbf{v} = A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

12. $T: M_{22} \to M_{22}$ definida mediante $T(A) = A - A^T$,

$\mathcal{B} = \mathcal{C} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$, $\mathbf{v} = A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

13. Considere el subespacio W de $\mathscr{D}$, dado por $W = \text{gen}(\text{sen } x, \cos x)$.

(a) Demuestre que el operador diferencial D mapea a W en sí mismo.

(b) Encuentre la matriz de D con respecto a $\mathcal{B} = \{\text{sen } x, \cos x\}$.

(c) Calcule indirectamente la derivada de $f(x) = 3 \text{ sen } x - 5 \cos x$, use el Teorema 6.26 y verifique que concuerda con $f'(x)$ como si se hubiera calculado directamente.

14. Considere el subespacio W de $\mathscr{D}$, dado por $W = \text{gen}(e^{2x}, e^{-2x})$.

(a) Demuestre que el operador diferencial D mapea a W en sí mismo.

(b) Encuentre la matriz de D con respecto a $\mathcal{B} = \{e^{2x}, e^{-x}\}$.

(c) Calcule indirectamente la derivada de $f(x) = e^{2x} - 3e^{-2x}$, use el Teorema 6.26 y verifique que concuerda con $f'(x)$ como si se hubiera calculado directamente.

15. Considere el subespacio W de $\mathscr{D}$, dado por $W = \text{gen}(e^{2x}, e^{2x} \cos x, e^{2x} \text{ sen } x)$.

(a) Encuentre la matriz de D con respecto a $\mathcal{B} = \{e^{2x}, e^{2x} \cos x, e^{2x} \text{ sen } x\}$.

(b) Calcule indirectamente la derivada de $f(x) = 3e^{2x} - e^{2x} \cos x + 2e^{2x} \text{ sen}$ usando el Teorema 6.26 y verifique que que concuerda con $f'(x)$ como si se hubiera calculado directamente.

16. Considere el subespacio W de $\mathscr{D}$, dado por $W = \text{gen}(\cos x, \text{sen } x, x \cos x, x \text{ sen } x)$.

(a) Encuentre la matriz de D con respecto a $\mathcal{B} = \{\cos x, \text{sen } x, x \cos x, x \text{ sen } x\}$.

(b) Calcule indirectamente la derivada de $f(x) = \cos x + 2x \cos x$, usando el Teorema 6.26 y verifique que concuerda con $f'(x)$ como si se hubiera calculado directamente.

En los ejercicios 17 y 18, $T: U \to V$ y $S: V \to W$ son transformaciones lineales, y $\mathcal{B}$, $\mathcal{C}$, y $\mathcal{D}$ son bases para U, V y W, respectivamente. Calcule $[S \circ T]_{\mathcal{D} \leftarrow \mathcal{B}}$ en dos formas: (a) al encontrar directamente $S \circ T$ y luego calcular su matriz y (b) al encontrar las matrices de S y T por separado y usar el Teorema 6.27.

17. $T: \mathscr{P}_1 \to \mathbb{R}^2$ definida mediante $T(p(x)) = \begin{bmatrix} p(0) \\ p(1) \end{bmatrix}$, $S: \mathbb{R}^2 \to \mathbb{R}^2$ definida mediante

$$S\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a-2b \\ 2a-b \end{bmatrix}, \mathcal{B} = \{1, x\}, \ \mathcal{C} = \mathcal{D} = \{\mathbf{e}_1, \mathbf{e}_2\}$$

18. $T: \mathscr{P}_1 \to \mathscr{P}_2$ definida mediante $T(p(x)) = p(x+1)$, $S: \mathscr{P}_2 \to \mathscr{P}_2$ definida mediante $S(p(x)) = p(x+1)$, $\mathcal{B} = \{1, x\}$, $\mathcal{C} = \mathcal{D} = \{1, x, x^2\}$

En los ejercicios 19-26, determine si la transformación lineal T es invertible al considerar su matriz con respecto a las bases estándar. Si T es invertible, use el Teorema 6.28 y el método del ejemplo 6.82 para encontrar T^{-1}.

19. T en el ejercicio 1

20. T en el ejercicio 5

21. T en el ejercicio 3

22. $T: \mathscr{P}_2 \to \mathscr{P}_2$ definida mediante $T(p(x)) = p'(x)$

23. $T : \mathscr{P}_2 \to \mathscr{P}_2$ definida mediante $T(p(x)) = p(x) + p'(x)$

24. $T : M_{22} \to M_{22}$ definida mediante $T(A) = AB$, donde

$$B = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$$

25. T en el ejercicio 11 **26.** T en el ejercicio 12

En los ejercicios 27-30, use el método del ejemplo 6.83 para evaluar la integral dada.

27. $\int (\operatorname{sen} x - 3\cos x)\,dx$. (Vea el ejercicio 13.)

28. $\int 5e^{-2x}\,dx$. (Vea el ejercicio 14.)

29. $\int (e^{2x}\cos x - 2e^{2x}\operatorname{sen} x)\,dx$. (Vea el ejercicio 15.)

30. $\int (x\cos x + x\operatorname{sen} x)\,dx$. (Vea el ejercicio 16.)

En los ejercicios 31-36, está dada una transformación lineal $T: V \to V$. Si es posible, encuentre una base $\mathcal{C}$ para V tal que la matriz $[T]_{\mathcal{C}}$ de T con respecto a $\mathcal{C}$ sea diagonal.

31. $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida mediante $T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} -4b \\ a + 5b \end{bmatrix}$

32. $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida mediante $T\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a - b \\ a + b \end{bmatrix}$

33. $T : \mathscr{P}_1 \to \mathscr{P}_1$ definida mediante $T(a + bx) = (4a + 2b) + (a + 3b)x$

34. $T : \mathscr{P}_2 \to \mathscr{P}_2$ definida mediante $T(p(x)) = p(x + 1)$

35. $T : \mathscr{P}_1 \to \mathscr{P}_1$ definida mediante $T(p(x)) = p(x) + xp'(x)$

36. $T : \mathscr{P}_2 \to \mathscr{P}_2$ definida mediante $T(p(x)) = p(3x + 2)$

37. Sea ℓ la recta que pasa por el origen en $\mathbb{R}^2$, con vector director $\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$. Use el método del ejemplo 6.85 para encontrar la matriz estándar de una reflexión en ℓ.

38. Sea W el plano en $\mathbb{R}^3$ con ecuación $x - y + 2z = 0$. Use el método del ejemplo 6.85 para encontrar la matriz estándar de una proyección ortogonal sobre W. Verifique que su respuesta es correcta usándola para calcular la proyección ortogonal de $\mathbf{v}$ sobre W, donde

$$\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$$

Compare su respuesta con el ejemplo 5.11. [*Sugerencia:* encuentre una descomposición ortogonal de $\mathbb{R}^3$ como $\mathbb{R}^3 = W + W^\perp$ usando una base ortogonal para W. Vea el ejemplo 5.3.]

39. Sea $T : V \to W$ una transformación lineal entre espacios vectoriales de dimensión finita y sean $\mathcal{B}$ y $\mathcal{C}$ bases para V y W, respectivamente. Demuestre que la matriz de T con respecto a $\mathcal{B}$ y $\mathcal{C}$ es única. Esto es, si A es una matriz tal que $A[\mathbf{v}]_{\mathcal{B}} = [T(\mathbf{v})]_{\mathcal{C}}$ para todo $\mathbf{v}$ en V, entonces $A = [T]_{\mathcal{C}\leftarrow\mathcal{B}}$. [*Sugerencia:* encuentre valores de $\mathbf{v}$ que demostrarán esto, una columna a la vez.]

En los ejercicios 40-45, sea $T : V \to W$ una transformación lineal entre espacios vectoriales de dimensión finita V y W. Sean $\mathcal{B}$ y $\mathcal{C}$ bases para V y W, respectivamente, y sea $A = [T]_{\mathcal{C}\leftarrow\mathcal{B}}$.

40. Demuestre que nulidad(T) = nulidad(A).

41. Demuestre que rank(T) = rank(A).

42. Si $V = W$ y $\mathcal{B} = \mathcal{C}$, demuestre que T es diagonalizable si y sólo si A es diagonalizable.

43. Use los resultados de esta sección para dar una prueba basada en las matrices del Teorema del rank (Teorema 6.19).

44. Si $\mathcal{B}'$ y $\mathcal{C}'$ también son bases para V y W, respectivamente, ¿cuál es la relación entre $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ y $[T]_{\mathcal{C}'\leftarrow\mathcal{B}'}$. Pruebe su afirmación.

45. Si dim $V = n$ y dim $W = m$, demuestre que $\mathscr{L}(V, W) \cong M_{mn}$. (Vea los ejercicios de la sección 6.4.) [*Sugerencia:* sean $\mathcal{B}$ y $\mathcal{C}$ bases para V y W, respectivamente. Demuestre que el mapeo $\varphi(T) = [T]_{\mathcal{C}\leftarrow\mathcal{B}}$, para T en $\mathscr{L}(V, W)$, define una transformación lineal φ: $\mathscr{L}(V, W) \to M_{mn}$ que es un isomorfismo.]

46. Si V es un espacio vectorial, entonces el ***espacio dual*** de V es el espacio vectorial $V^* = \mathscr{L}(V, \mathbb{R})$. Demuestre que si V es de dimensión finita, entonces $V^* \cong V$.

Exploración

Mosaicos, retículas y la restricción cristalográfica

En la naturaleza y el arte frecuentemente se encuentran patrones. La estructura molecular de los cristales con frecuencia muestran repetición, así como los azulejos y mosaicos que se encuentran en las obras artísticas de muchas culturas. *Embaldosar* (o *azulejar*) es cubrir un plano con formas que no se traslapan ni dejan espacios. El artista holandés M. C. Escher (1898-1972) produjo muchas obras en las que exploró la posibilidad de embaldosar un plano con figuras fantásticas (figura 6.14).



Figura 6.14
"Dibujo simétrico E103" de M. C. Escher

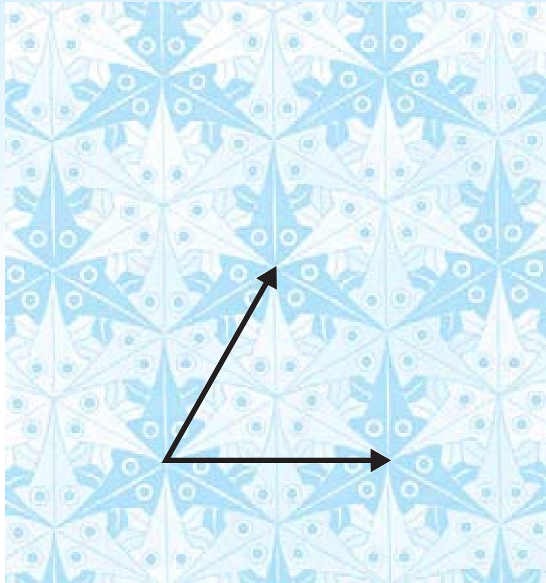


Figura 6.15
Invarianza bajo traslación
"Dibujo simétrico E103" de M. C. Escher

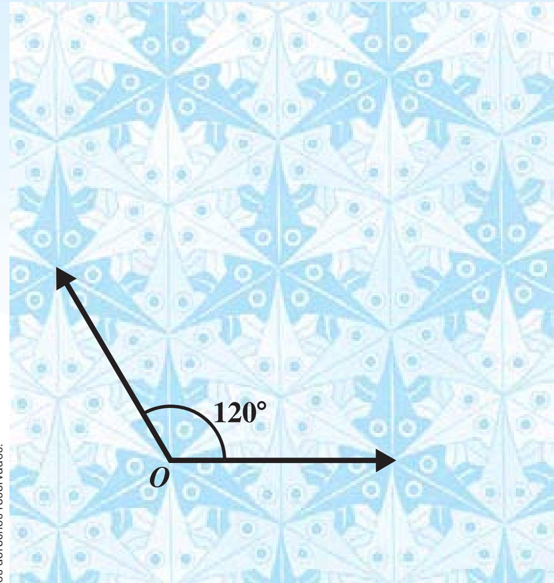


Figura 6.17
Simetría rotacional
"Dibujo simétrico E103" de M. C. Escher

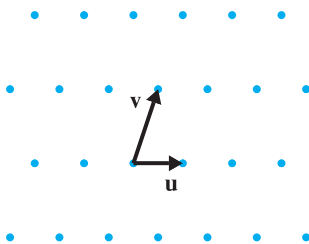


Figura 6.16
Una retícula

Esta exploración se interesa en los patrones como los de la figura 6.14, que se supone son infinitos y repetitivos en todas las direcciones del plano. Tal patrón tiene la propiedad de que puede recorrerse (o *trasladarse*) en al menos dos direcciones (que corresponden a dos vectores linealmente independientes), de modo que parece que no se movieron en absoluto. Se dice que el patrón es *invariante* bajo traslaciones y tiene ***simetría traslacional*** en dichas direcciones. Por ejemplo, el patrón de la figura 6.14 tiene simetría traslacional en las direcciones que se muestran en la figura 6.15.

Si un patrón tiene simetría traslacional en dos direcciones, tiene simetría traslacional en infinitas direcciones.

1. Sean **u** y **v** los dos vectores que se muestran en la figura 6.15. Demuestre que el patrón de la figura 6.14 es invariante por traslación para cualquier combinación lineal *entera* de **u** y **v**; esto es, por cualquier vector de la forma $a\mathbf{u} + b\mathbf{v}$, donde a y b son enteros.

Para cualesquiera vectores linealmente independientes **u** y **v** en $\mathbb{R}^2$, el conjunto de puntos determinado por todas las combinaciones lineales enteras de **u** y **v** se llama ***retícula***. La figura 6.16 muestra un ejemplo de una retícula.

2. Dibuje la retícula correspondiente a los vectores **u** y **v** de la figura 6.15.

La figura 6.14 también muestra ***simetría rotacional***. Esto es, es posible rotar todo el patrón en torno a algún punto y que parezca no cambiar. Se dice que es *invariante* ante tal rotación. Por ejemplo, el patrón de la figura 6.14 es invariante ante una rotación de 120° en torno al punto O, como se muestra en la figura 6.17. A O se le llama ***centro*** de simetría rotacional (o ***centro de rotación***).

Note que si un patrón se basa en una retícula subyacente, entonces la retícula también debe poseer cualquier simetría del patrón.

3. Explique por qué si un punto O es un centro de rotación en un ángulo θ, entonces es un centro de rotación para todo múltiplo entero de θ. Deduzca que, si $0 < \theta \leq 360°$, entonces $360/\theta$ debe ser un entero. (Si $360/\theta = n$, se dice que el patrón o retícula tiene simetría rotacional ***n*-tuple**.)

4. ¿Cuál es el menor ángulo positivo de simetría rotacional para la retícula del problema 2? ¿El patrón de la figura 6.14 también tiene simetría rotacional a través de este ángulo?

5. Tome varios valores de θ tales que $0 < \theta \leq 360°$ y $360/\theta$ es entero. Intente dibujar una retícula que tenga simetría rotacional a través del ángulo θ. En particular, ¿puede dibujar una retícula con simetría rotacional óctuple?

Se demostrará que los valores de θ que son posibles ángulos de simetría rotacional para una retícula están severamente restringidos. La técnica que se usará es considerar transformaciones rotacionales en términos de diferentes bases. En concordancia, sea R_θ una rotación en torno al origen en un ángulo θ y sea $\mathcal{E}$ la base estándar para $\mathbb{R}^2$. Entonces la matriz estándar de R_θ es

$$[R_\theta]_\mathcal{E} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

6. Con referencia a los problemas 2 y 4, tome el origen como los orígenes de **u** y **v**.

(a) ¿Cuál es el valor real (es decir, numérico) de $[R_\theta]_\mathcal{E}$ en este caso?
(b) Sea $\mathcal{B}$ la base $\{\mathbf{u}, \mathbf{v}\}$. Calcule la matriz $[R_\theta]_\mathcal{B}$.

7. En general, sean **u** y **v** dos vectores linealmente independientes cualesquiera en $\mathbb{R}^2$ y suponga que la retícula determinada por **u** y **v** es invariante ante una rotación en un ángulo θ. Si $\mathcal{B} = \{\mathbf{u}, \mathbf{v}\}$, demuestre que la matriz de R_θ con respecto a $\mathcal{B}$ debe tener la forma

$$[R_\theta]_\mathcal{B} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

donde a, b, c y d son *enteros*.

8. En la terminología y notación del problema 7, demuestre que $2\cos\theta$ debe ser un entero. [*Sugerencia:* use el ejercicio 35 de la sección 4.4 y el Teorema 6.29.]

9. Con el problema 8, elabore una lista de todos los posibles valores de θ, con $0 < \theta \leq 360°$, que puedan ser ángulos de simetría rotacional de una retícula. Registre los valores correspondientes de n, donde $n = 360/\theta$, para demostrar que una retícula puede tener simetría rotacional n-tuple si y sólo si $n = 1, 2, 3, 4$ o 6. Este resultado, conocido como la ***restricción cristalográfica***, la probó por primera vez W. Barlow en 1894.

10. En la biblioteca o la Internet, vea si puede encontrar un mosaico de Escher para cada uno de los cinco tipos posibles de simetría rotacional; esto es, donde el *menor* ángulo de simetría rotacional del patrón sea uno de los especificados por la restricción cristalográfica.

6.7 Aplicaciones

Ecuaciones diferenciales lineales homogéneas



En los ejercicios 69-72 de la sección 4.6, se demostró que si $y = y(t)$ es una función doblemente derivable que satisface la ecuación diferencial

$$y'' + ay' + by = 0 \tag{1}$$

entonces y es de la forma

$$y = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

si λ_1 y λ_2 son *distintas* raíces de la ecuación característica asociada $\lambda^2 + a\lambda + b = 0$. (El caso en que $\lambda_1 = \lambda_2$ se dejó sin resolver.) El ejemplo 6.12 y el ejercicio 20 de esta sección demuestran que el conjunto de soluciones a la ecuación (1) forman un subespacio de $\mathscr{F}$, el espacio vectorial de funciones. En esta sección se da mayor seguimiento a estas ideas, y se pone particular atención en el papel que desempañan los espacios vectoriales, las bases y la dimensión.

Para preparar el escenario, considere una clase más simple de ejemplos. Una ecuación diferencial de la forma

$$y' + ay = 0 \tag{2}$$

se llama ***ecuación diferencial lineal homogénea de primer orden***. ("Primer orden" se refiere al hecho de que la derivada más alta que se involucra es una primera derivada y "homogénea" significa que el lado derecho es cero. ¿Ve por qué la ecuación es "lineal"?) Una ***solución*** a la ecuación (2) es una función derivable $y = y(t)$ que satisface la ecuación (2) para todos los valores de t.



Es fácil comprobar que una solución a la ecuación (2) es $y = e^{-at}$. (Hágalo.) Sin embargo, se quieren describir *todas* las soluciones, y es aquí donde entran los espacios vectoriales. Se tiene el siguiente teorema.

Teorema 6.31 El conjunto S de todas las soluciones de $y' + ay = 0$ es un subespacio de $\mathscr{F}$.

Demostración Dado que la función cero ciertamente satisface la ecuación (2), S no está en el vacío. Sean x y y dos funciones derivables de t que están en S y sea c un escalar. Entonces

$$x' + ax = 0 \quad \text{y} \quad y' + ay = 0$$

de modo, que al usar las reglas de derivación, se tiene

$$(x + y)' + a(x + y) = x' + y' + ax + ay = (x' + ax) + (y' + ay) = 0 + 0 = 0$$

y

$$(cy)' + a(cy) = cy' + c(ay) = c(y' + ay) = c \cdot 0 = 0$$

Por tanto, $x + y$ y cy también están en S, de modo que S es un subespacio de $\mathscr{F}$.

Ahora se demostrará que S es un subespacio unidimensional de $\mathscr{F}$ y que $\{e^{-at}\}$ es una base. Para este fin, sea $x = x(t)$ que está en S. Entonces, para toda t,

$$x'(t) + ax(t) = 0 \quad \text{o} \quad x'(t) = -ax(t)$$

Defina una nueva función $z(t) = x(t)e^{at}$. Luego, por la regla de la cadena para la derivación,

$$\begin{aligned} z'(t) &= x(t)ae^{at} + x'(t)e^{at} \\ &= ax(t)e^{at} - ax(t)e^{at} \\ &= 0 \end{aligned}$$

Dado que z' es idénticamente cero, z debe ser una función constante, por decir, $z(t) = k$. Pero esto significa que

$$x(t)e^{at} = z(t) = k \quad \text{para toda } t$$

de modo que $x(t) = ke^{-at}$. Por tanto, todas las soluciones a la ecuación (2) son múltiplos escalares de la solución única $y = e^{-at}$. Se demostró el siguiente teorema.

Teorema 6.32 Si S es el espacio solución de $y' + ay = 0$, entonces $\dim S = 1$ y $\{e^{-at}\}$ es una base para S.

Un modelo para el crecimiento poblacional supone que la tasa de crecimiento de la población es proporcional al tamaño de la población. Este modelo funciona bien si existen pocas restricciones (como espacio o alimento limitados, o cosas por el estilo) para el crecimiento. Si el tamaño de la población en el tiempo t es $p(t)$, entonces la tasa de crecimiento, o razón de cambio de la población, es su derivada $p'(t)$. La suposición de que la tasa de crecimiento de la población es proporcional a su tamaño puede escribirse como

$$p'(t) = kp(t)$$

donde k es la constante de proporcionalidad. Por ende, p satisface la ecuación diferencial $p' - kp = 0$, de modo que, por el Teorema 6.32,

$$p(t) = ce^{kt}$$

para algún escalar c. Las constantes c y k se determinan con el uso de datos experimentales.

Ejemplo 6.87 La bacteria *Escherichia coli* (o *E. coli*, para abreviar) se encuentra comúnmente en los intestinos de los humanos y otros mamíferos. Plantea severos riesgos a la salud si escapa hacia el ambiente. En condiciones de laboratorio, cada célula de la bacteria se divide en dos cada 20 minutos. Si comienza con una sola célula de *E. coli*, ¿cuántas habrá después de un día?

La *E. coli* se menciona en la novela de Michael Crichton, *The Andromeda Strain* (Nueva York: Dell, 1969), aunque el "villano" en la novela supuestamente fue un virus extraterrestre. En la vida real, la *E. coli* contaminó el suministro de agua de la ciudad de Walkerton, Ontario, en 2000, lo que resultó en siete muertes y provocó que cientos de personas enfermaran seriamente.

Solución No es necesario usar ecuaciones diferenciales para resolver este problema, pero se hará, con la finalidad de ilustrar el método básico.

Para determinar c y k, se usan los datos dados en el enunciado del problema. Si se toma 1 unidad de tiempo como 20 minutos, entonces se tiene $p(0) = 1$ y $p(1) = 2$. Por tanto,

$$c = c \cdot 1 = ce^{k \cdot 0} = 1 \quad \text{y} \quad 2 = ce^{k \cdot 1} = e^k$$

Se tiene que $k = \ln 2$, de modo que

$$p(t) = e^{t \ln 2} = e^{\ln 2^t} = 2^t$$

Después de 1 día, $t = 72$, de modo que el número de células bacterianas será $p(72) = 2^{72} \approx 4.72 \times 10^{21}$ (vea la figura 6.18).

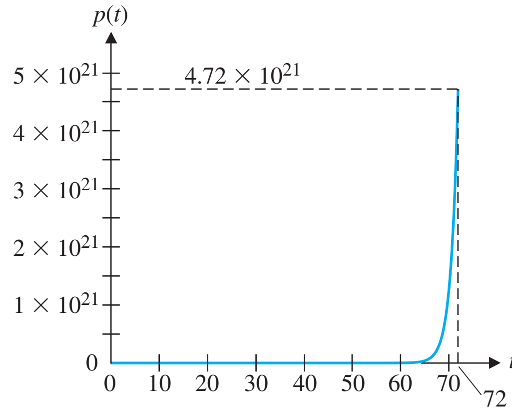
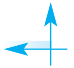


Figura 6.18
Crecimiento exponencial



Las sustancias radiactivas se desintegran al emitir radiación. Si $m(t)$ denota la masa de la sustancia en el tiempo t, entonces la tasa de desintegración es $m'(t)$. Los físicos descubrieron que la tasa de desintegración de una sustancia es proporcional a su masa; esto es,

$$m'(t) = km(t) \quad \text{o} \quad m' - km = 0$$

donde k es una constante negativa. Al aplicar el Teorema 6.32, se tiene

$$m(t) = ce^{kt}$$

para alguna constante c. El tiempo requerido para que la mitad de una sustancia radiactiva se desintegre se llama ***vida media***.

Ejemplo 6.88

Después de 5.5 días, una muestra de 100 mg de radón 222 desintegra a 37 mg.

(a) Encuentre una fórmula para $m(t)$, la masa restante después de t días.

(b) ¿Cuál es la vida media del radón 222?

(c) ¿Cuándo quedarán solamente 10 mg?

Solución (a) A partir de $m(t) = ce^{kt}$, se tiene

$$100 = m(0) = ce^{k \cdot 0} = c \cdot 1 = c$$

de modo que $$m(t) = 100e^{kt}$$

Con el tiempo medido en días, se tiene $m(5.5) = 37$. Por tanto,

$$100e^{5.5k} = 37$$

de modo que $$e^{5.5k} = 0.37$$

Al despejar k, se encuentra

$$5.5k = \ln(0.37)$$

de modo que $$k = \frac{\ln(0.37)}{5.5} \approx -0.18$$

Por tanto, $m(t) = 100e^{-0.18t}$.

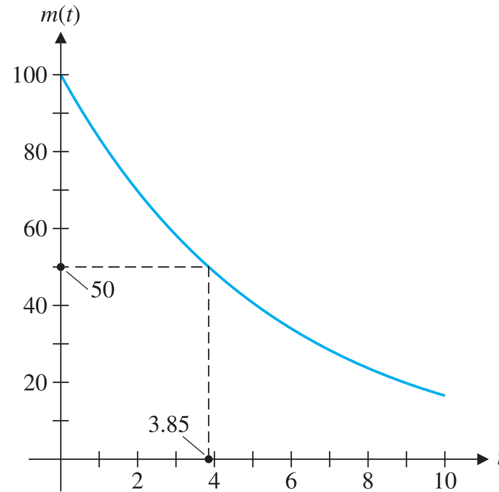


Figura 6.19
Decaimiento radioactivo

(b) Para encontrar la vida media del radón 222, se necesita el valor de t para el cual $m(t) = 50$. Al resolver esta ecuación, se encuentra

$$100e^{-0.18t} = 50$$

de modo que

$$e^{-0.18t} = 0.50$$

Por tanto,

$$-0.18t = \ln(\tfrac{1}{2}) = -\ln 2$$

y

$$t = \frac{\ln 2}{0.18} \approx 3.85$$

Por ende, el radón 222 tiene una vida media de aproximadamente 3.85 días. (Vea la figura 6.19.)

(c) Es necesario determinar el valor de t tal que $m(t) = 10$. Esto es, se debe resolver la ecuación

$$100e^{-0.18t} = 10 \quad \text{o} \quad e^{-0.18t} = 0.1$$

Al tomar el logaritmo natural de ambos lados produce $-0.18t = \ln 0.1$. Por tanto,

$$t = \frac{\ln 0.1}{-0.18} \approx 12.79$$

de modo que, después de aproximadamente 12.79 días, quedarán 10 mg de la muestra.

El conjunto solución de la ecuación diferencial de segundo orden $y'' + ay' + by = 0$ también es un subespacio de $\mathscr{F}$ (ejercicio 20) y se evidencia que la dimensión de S es 2. El inciso (a) del Teorema 6.33, que se extiende al Teorema 6.32, se implica con el Teorema 4.40. Aquí el planteamiento es usar el poder de los espacios vectoriales; hacerlo permite obtener también la parte (b) del Teorema 6.33, un resultado que no podría obtenerse con los métodos previos.

Vea *Linear Algebra* de S. H. Friedberg, A. J. Insel y L. E. Spence (Englewood Cliffs, NJ: Prentice-Hall, 1979).

Teorema 6.33 Sea S el espacio solución de

$$y'' + ay' + by = 0$$

y sean λ_1 y λ_2 las raíces de la ecuación característica $\lambda^2 + a\lambda + b = 0$.

a. Si $\lambda_1 \neq \lambda_2$, entonces $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ es una base para S.

b. Si $\lambda_1 = \lambda_2$, entonces $\{e^{\lambda_1 t}, te^{\lambda_1 t}\}$ es una base para S.

Comentarios

- Observe que lo que dice el teorema, en otras palabras, es que las soluciones de $y'' + ay' + by = 0$ son de la forma

$$y = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

en el primer caso y

$$y = c_1 e^{\lambda_1 t} + c_2 t e^{\lambda_1 t}$$

en el segundo caso.

- Compare el Teorema 6.33 con el Teorema 4.38. Las ecuaciones diferenciales lineales y las relaciones de recurrencia lineal tienen mucho en común. Aunque las primeras pertenecen a las matemáticas *continuas* y las últimas a las matemáticas *discretas*, existen muchos paralelismos.

Demostración (a) Demuestre primero que $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ está contenido en S. Sea λ cualquier raíz de la ecuación característica y sea $f(t) = e^{\lambda t}$. Entonces

$$f'(t) = \lambda e^{\lambda t} \quad \text{y} \quad f''(t) = \lambda^2 e^{\lambda t}$$

de donde se tiene que

$$\begin{aligned} f'' + af' + bf &= \lambda^2 e^{\lambda t} + a\lambda e^{\lambda t} + be^{\lambda t} \\ &= (\lambda^2 + a\lambda + b)e^{\lambda t} \\ &= 0 \cdot e^{\lambda t} = 0 \end{aligned}$$

Por tanto, f está en S. Pero, dado que λ_1 y λ_2 son raíces de la ecuación característica, esto significa que $e^{\lambda_1 t}$ y $e^{\lambda_2 t}$ están en S.

El conjunto $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ también es linealmente independiente, porque si

$$c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} = 0$$

entonces, al hacer $t = 0$, se tiene

$$c_1 + c_2 = 0 \quad \text{o} \quad c_2 = -c_1$$

A continuación, se hace $t = 1$ para obtener

$$c_1 e^{\lambda_1} - c_1 e^{\lambda_2} = 0 \quad \text{o} \quad c_1(e^{\lambda_1} - e^{\lambda_2}) = 0$$

Pero, $e^{\lambda_1} - e^{\lambda_2} \neq 0$, pues $e^{\lambda_1} - e^{\lambda_2} = 0$ implica que $e^{\lambda_1} = e^{\lambda_2}$ lo que claramente es imposible si $\lambda_1 \neq \lambda_2$. (Vea la figura 6.20.) Se deduce que $c_1 = 0$ y, por tanto, $c_2 = 0$, de modo que $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ es linealmente independiente.

Dado que $S = 2$, $\{e^{\lambda_1 t}, e^{\lambda_2 t}\}$ debe ser una base para S.

(b) Se le pide demostrar esta propiedad en el ejercicio 21.

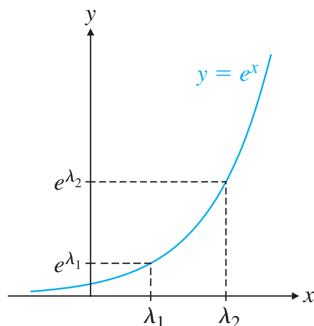


Figura 6.20

Ejemplo 6.89

Encuentre todas las soluciones de $y'' - 5y' + 6y = 0$.

Solución La ecuación característica es $\lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3) = 0$. Por tanto, las raíces son 2 y 3, de modo que $\{e^{2t}, e^{3t}\}$ es una base para el espacio solución. Se tiene que las soluciones a la ecuación dada son de la forma

$$y = c_1e^{2t} + c_2e^{3t}$$

Las constantes c_1 y c_2 pueden determinarse si se especifican ecuaciones adicionales, llamadas ***condiciones de frontera***.

Ejemplo 6.90

Encuentre la solución de $y'' + 6y' + 9y = 0$ que satisfaga $y(0) = 1, y'(0) = 0$.

Solución La ecuación característica es $\lambda^2 + 6\lambda + 9 = (\lambda + 3)^2 = 0$, de modo que -3 es una raíz repetida. Por tanto, $\{e^{-3t}, te^{-3t}\}$ es una base para el espacio solución y la solución general es de la forma

$$y = c_1e^{-3t} + c_2te^{-3t}$$

La primera condición de frontera produce

$$1 = y(0) = c_1e^{-3\cdot 0} + 0 = c_1$$

de modo que $y = e^{-3t} + c_2te^{-3t}$. Al derivar, se tiene

$$y' = -3e^{-3t} + c_2(-3te^{-3t} + e^{-3t})$$

así que la segunda condición de frontera produce

$$0 = y'(0) = -3e^{-3\cdot 0} + c_2(0 + e^{-3\cdot 0}) = -3 + c_2$$

o $$c_2 = 3$$

Por tanto, la solución requerida es

$$y = e^{-3t} + 3te^{-3t} = (1 + 3t)e^{-3t}$$

a + bi

El Teorema 6.33 incluye el caso en el que las raíces de la ecuación característica son complejas. Si $\lambda = p + qi$ es una raíz compleja de la ecuación $\lambda^2 + a\lambda + b = 0$, entonces también lo es su conjugado $\bar{\lambda} = p - qi$. (Vea los Apéndices C y D.) Por el Teorema 6.33(a), el espacio solución S de la ecuación diferencial $y'' + ay' + by = 0$ tiene $\{e^{\lambda t}, e^{\bar{\lambda} t}\}$ como una base. Ahora,

$$e^{\lambda t} = e^{(p+qi)t} = e^{pt}e^{i(qt)} = e^{pt}(\cos qt + i \operatorname{sen} qt)$$

y $$e^{\bar{\lambda} t} = e^{(p-qi)t} = e^{pt}e^{i(-qt)} = e^{pt}(\cos qt - i \operatorname{sen} qt)$$

de modo que $$e^{pt}\cos qt = \frac{e^{\lambda t} + e^{\bar{\lambda} t}}{2} \quad \text{y} \quad e^{pt}\operatorname{sen} qt = \frac{e^{\lambda t} - e^{\bar{\lambda} t}}{2i}$$

Se tiene que $\{e^{pt}\cos qt, e^{pt}\operatorname{sen} qt\}$ está contenido en gen $(e^{\lambda t}, e^{\bar{\lambda} t}) = S$. Puesto que $e^{pt}\cos qt$ y $e^{pt}\operatorname{sen} qt$ son linealmente independientes (vea el ejercicio 22) y dim $S = 2$, $\{e^{pt}\cos qt, e^{pt}\operatorname{sen} qt\}$ también es una base para S. Por tanto, cuando su ecuación característica tiene una raíz compleja $p + qi$, la ecuación diferencial $y'' + ay' + by = 0$ tiene soluciones de la forma

$$y = c_1e^{pt}\cos qt + c_2e^{pt}\operatorname{sen} qt$$

a + bi **Ejemplo 6.91**

Encuentre todas las soluciones de $y'' - 2y' + 4 = 0$.

Solución La ecuación característica es $\lambda^2 - 2\lambda + 4 = 0$ con raíces $1 \pm i\sqrt{3}$. El análisis anterior dice que la solución general a la ecuación diferencial dada es

$$y = c_1e^t \cos \sqrt{3}t + c_2e^t \operatorname{sen} \sqrt{3}t$$

a + bi **Ejemplo 6.92**

Una masa está unida al extremo de un resorte vertical (figura 6.21). Si la masa se jala hacia abajo y se libera, oscilará arriba y abajo. Dos leyes de la física gobiernan esta situación. La primera, la ***ley de Hooke***, afirma que si el resorte se estira (o comprime) x unidades, la fuerza F necesaria para restaurarlo a su posición original es proporcional a x:

$$F = -kx$$

donde k es una constante positiva (llamada constante del resorte). La ***segunda ley del movimiento de Newton*** afirma que la fuerza es igual a la masa por la aceleración. Dado que $x = x(t)$ representa la distancia, o desplazamiento, del resorte en el tiempo t, x' proporciona su velocidad y x'' su aceleración. Por tanto, se tiene

$$mx'' = -kx \quad \text{o} \quad x'' + \left(\frac{k}{m}\right)x = 0$$

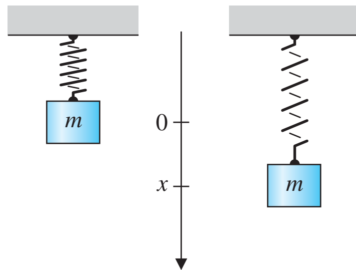


Figura 6.21

Puesto que tanto k como m son positivas, también lo es $K = k/m$ y la ecuación diferencial tiene la forma $x'' + Kx = 0$, donde K es positiva.

La ecuación característica es $\lambda^2 + K = 0$, con raíces $\pm i\sqrt{K}$. Por tanto, la solución general a la ecuación diferencial del resorte en oscilación es

$$x = c_1 \cos \sqrt{K}t + c_2 \operatorname{sen} \sqrt{K}t$$

Suponga que el resorte está en reposo ($x = 0$) en el tiempo $t = 0$ segundos y se estira tanto como sea posible, hasta una longitud de 20 cm, antes de soltarse. Entonces

$$0 = x(0) = c_1 \cos 0 + c_2 \operatorname{sen} 0 = c_1$$

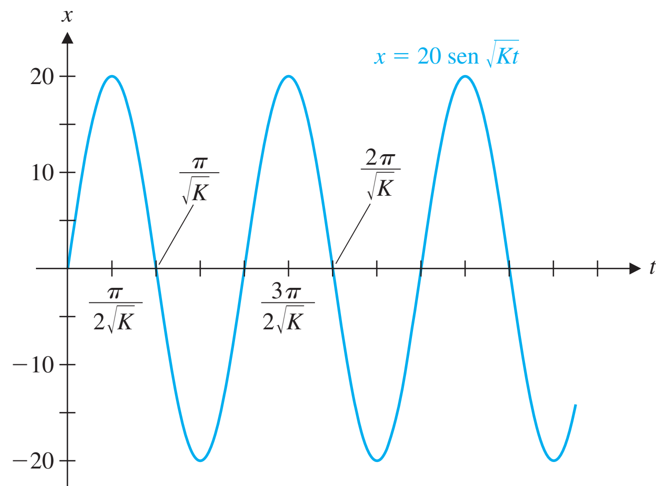


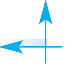
Figura 6.22

de modo que $x = c_2 \operatorname{sen} \sqrt{K}t$. Puesto que el valor máximo de la función seno es 1, debe tener $c_2 = 20$ (lo que ocurre por primera vez cuando $t = \pi/2\sqrt{K}$), lo que produce la solución

$$x = 20 \operatorname{sen} \sqrt{K}t$$

(Vea la figura 6.22.)

Desde luego, esta es una solución idealizada, pues ignora cualquier forma de resistencia y predice que el resorte oscilará por siempre. Es posible tomar en cuenta efectos de amortiguamiento (como la fricción), pero este modelo simple sirvió para introducir una importante aplicación de las ecuaciones diferenciales y las técnicas que aquí se desarrollaron.



Códigos lineales

Ahora se dirigirá la atención hacia la clase de códigos más importantes y más ampliamente usados: los *códigos lineales*. De hecho, muchos de los ejemplos ya analizados caen en esta categoría. La NASA usa de manera extensa los códigos lineales para transmitir imágenes del espacio exterior. En 1972, la nave espacial *Mariner 9* usó un tipo de código lineal llamado *código de Reed-Muller* para transmitir imágenes en blanco y negro de Marte (figura 6.23). Luego, entre 1979 y 1981, el *Voyager 1* y el *Voyager 2* pudieron enviar de vuelta notables imágenes a color de Júpiter y Saturno (que se reproducen en blanco y negro en la figura 6.24) usando un *código Golay*, otro código lineal.

Definición Un ***código lineal*** p-ario es un subespacio C de $\mathbb{Z}_p^n$.

Como siempre, el principal interés es el caso $p = 2$, los ***códigos lineales binarios***. Comprobar si un subconjunto C de $\mathbb{Z}_2^n$ es un subespacio involucra demostrar que C satisface las condiciones del Teorema 6.32. Dado que en $\mathbb{Z}_2^n$ los únicos escalares son 0 y 1, verificar que C es cerrado para la multiplicación por un escalar sólo involucra demostrar que C contiene el vector cero. Todo lo que resta por comprobar es que C sea cerrado para la suma.


NASA-JPL

Figura 6.23
El polo sur de Marte


NASA-JPL

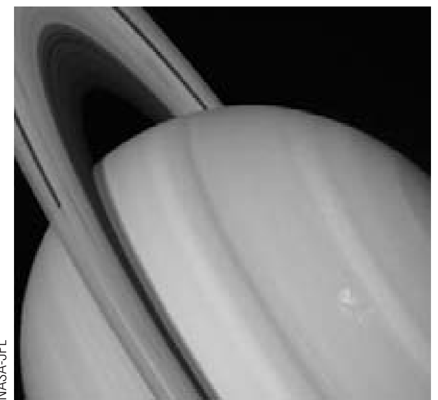

NASA-JPL

Figura 6.24
La mancha roja de Júpiter y los anillos de Saturno

Ejemplo 6.93

¿$C_1 = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$ y $C_2 = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ son códigos lineales (binarios)?

Solución Claramente, C_1 contiene el vector cero y es cerrado para la suma, de modo que es un código lineal. C_2 no es cerrado para la suma, pues no contiene

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Por tanto, C_2 no es lineal.

Para el resto de la sección, se omitirá el adjetivo "binario", pues todos los códigos que se considerarán serán binarios. Si un código lineal C es un subespacio k-dimensional de $\mathbb{Z}_2^n$, entonces se dice que C es un ***código (n, k)***.

Ejemplo 6.94

(a) El código C_1 del ejemplo 6.93 es un subespacio de $\mathbb{Z}_2^4$ y tiene dimensión 2, pues

$$\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

es una base para C_1. (De hecho, C_1 tiene exactamente tres bases diferentes de dos elementos. ¿Cuáles son las otras dos? Vea el ejercicio 31.) En consecuencia, C_1 es un código (4, 2).

(b) El código H de Hamming (7, 4), que se introdujo en la sección 3.7, es un código lineal (7, 4) (¡por fortuna!), en la nueva terminología. Es lineal porque tiene una matriz generadora G, de modo que sus vectores son todos los vectores de la forma $G\mathbf{x}$ donde $\mathbf{x}$ está en $\mathbb{Z}_2^4$. Pero este es justo el espacio columna de la matriz G de 7×4 y por tanto es un subespacio de $\mathbb{Z}_2^7$. Puesto que las cuatro columnas de G son linealmente independientes (¿por qué?), forman una base para H. Por tanto, H es un código (7, 4).

(c) Los códigos

$$C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \quad \text{y} \quad C^{\perp} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

son códigos duales. Es fácil ver que cada uno de ellos es un código lineal, que $\dim C = 1$, y que $\dim C^{\perp} = 2$. (Compruebe esta afirmación.) Por tanto, C es un código (3, 1) y $C^{\perp}$ es un código (3, 2). El hecho de que $3 = 1 + 2$ no es un accidente, como muestra el siguiente teorema.

Teorema 6.34

Sea C un código lineal (n, k).

a. El código dual $C^{\perp}$ es un código lineal $(n, n - k)$.
b. C contiene 2^k vectores, y $C^{\perp}$ contiene 2^{n-k} vectores.

Demostración (a) Puesto que C es un código lineal (n, k), es un subespacio k-dimensional de $\mathbb{Z}_2^n$. Su dual $C^\perp$ es el complemento ortogonal de C y por tanto también es un subespacio de $\mathbb{Z}_2^n$, por el Teorema 5.9(a). En consecuencia, $C^\perp$ es un código lineal.

Ahora es posible aplicar el Teorema 5.13 para demostrar que

$$\dim C^\perp = n - \dim C = n - k$$



(*Nota:* los Teoremas 5.9 y 5.13 son verdaderos si $\mathbb{R}^n$ se sustituye con $\mathbb{Z}_2^n$. Éste es el caso para la mayoría de los resultados no geométricos acerca de la ortogonalidad.) Se concluye que $C^\perp$ es un código$(n, n-k)$.

(b) Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ una base para C. Entonces los vectores en C son todos los vectores de la forma

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$$

donde cada c_i es 0 o 1. Por tanto, existen dos posibilidades para c_1 y, para cada una de ellas, dos posibilidades de c_2, y así sucesivamente, lo que hace el número total de posibilidades para $\mathbf{v}$

$$\underbrace{2 \times 2 \times \cdots \times 2}_{k \text{ veces}} = 2^k$$

Por tanto, C contiene exactamente 2^k vectores. Al aplicar esta fórmula a su código dual $(n, n-k)$, se ve que $C^\perp$ tiene 2^{n-k} vectores.

Los códigos de Reed-Muller reciben su nombre en honor de los científicos de computación Irving S. Reed y David E. Muller, quienes en 1954 publicaron artículos, de manera independiente, acerca de esos códigos.

Recuerde que la representación binaria, o de base dos, de un número surge de escribirlo como una suma de distintas potencias de dos. Si $n = b_k \cdot 2^k + \cdots + b_1 \cdot 2 + b_0$, donde cada b_i es 0 o 1, entonces en base dos n se representa como $n = b_i \cdots b_1 b_0$. Por ejemplo, $25 = 16 + 8 + 1 = 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2 + 1$, de modo que la representación binaria de 25 es 11001.

Ahora se construye una de las familias más antiguas de códigos lineales, los códigos de Reed-Muller. Como se mencionó anteriormente, este es el tipo de código que usó la nave espacial *Mariner 9* para transmitir imágenes de Marte. Con la finalidad de transmitirse, cada fotografía tiene que descomponerse en elementos de imágenes, o *píxeles*. Esto se hizo al suponer a la fotografía una rejilla de 700×832 píxeles y luego asignar a cada píxel uno de 64 tonos de gris, que varían de blanco (0) a negro (63). Puesto que $64 = 2^6$, puede usarse aritmética binaria para representar cada una de dichas sombras: blanco es 000000 y negro es 111111. Entonces puede reescribir estos 64 números binarios como vectores en $\mathbb{Z}_2^6$ y codificarlos usando un código que corrige tantos errores como sea posible. El código que se eligió usar en el *Mariner 9* pertenece a una gran familia de códigos que se definen más fácilmente de manera inductiva.

Definición Los ***códigos de Reed-Muller*** R_n (de primer orden) se definen inductivamente del modo siguiente:

1. Para $n = 0$, $R_0 = \mathbb{Z}_2 = \{0, 1\}$.
2. Para $n \geq 1$, R_n es el subespacio de $\mathbb{Z}_2^{2^n}$ cuya base consiste de todos los vectores de la forma

$$\begin{bmatrix} \mathbf{u} \\ \mathbf{u} \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}$$

donde $\mathbf{u}$ es un vector base en R_{n-1}, $\mathbf{0}$ es el vector cero en $\mathbb{Z}_2^{2^{n-1}}$, y $\mathbf{1}$ es el vector de números 1 en $\mathbb{Z}_2^{2^{n-1}}$.

Para tener un sentido de lo que contienen estos vectores, use la definición para construir R_1 y R_2. Una base para $R_0 = \mathbb{Z}_2$ es justo $\{1\}$, de modo que una base para R_1 es

$$\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$$

Por tanto, por cerradura bajo la suma, R_1 también debe contener a los vectores

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Es fácil comprobar que ningún otro vector puede obtenerse mediante la suma, de modo que

$$R_1 = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} = \mathbb{Z}_2^2$$

De igual modo, una base para R_2 es

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} \right\}$$

y, por cerradura bajo la suma, es fácil comprobar que los $8 = 2^3$ vectores en R_2 son

$$R_2 = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Note que, en R_1, todo vector código, excepto $\mathbf{0}$ y $\mathbf{1}$, tiene peso 1, y en R_2 todo vector código, excepto $\mathbf{0}$ y $\mathbf{1}$, tiene peso 2. Esta es una propiedad general de los códigos de Reed-Muller, y se prueba como parte del siguiente teorema. Pero primero note que el ***complemento*** de un vector $\mathbf{x}$ en $\mathbb{Z}_2^n$ es el vector $\overline{\mathbf{x}}$ que se obtiene al cambiar todos los ceros por 1 y viceversa. Por ejemplo,

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix} \Leftrightarrow \overline{\mathbf{x}} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

Observe que $\overline{\mathbf{x}} = \mathbf{x} + \mathbf{1}$, donde $\mathbf{1}$ es el vector que consiste por completo de números 1.

Teorema 6.35 Para $n \geq 1$, el código de Reed-Muller R_n es un código lineal $(2^n, n + 1)$ en el que todo vector código, excepto $\mathbf{0}$ y $\mathbf{1}$, tiene peso 2^{n-1}.

Demostración Este teorema se demostrará por inducción sobre n. Para $n = 1$, ya se vio que $R_1 = \mathbb{Z}_2^2$ es un código lineal $(2, 2) = (2^1, 1 + 1)$ en el que todo vector código, excepto $\mathbf{0}$ y $\mathbf{1}$, tiene peso $1 = 2^{1-1}$. Suponga que el resultado es verdadero para $n = k$; esto es, suponga que R_k es un código lineal $(2^k, k + 1)$ en el que todo vector código, excepto $\mathbf{0}$ y $\mathbf{1}$, tiene peso 2^{k-1}. Ahora considere R_{k+1}.

Por construcción, R_{k+1} tiene una base que consiste de vectores de la forma $\begin{bmatrix} \mathbf{u} \\ \mathbf{u} \end{bmatrix}$, donde **u** está en R_k, junto con el vector $\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}$. Por la hipótesis de inducción, los vectores **u**, **0** y **1** están en $\mathbb{Z}_2^{2^k}$; por tanto, los vectores base para R_{k+1} están en $\mathbb{Z}_2^{2^{k+1}}$. Más aún, la dimensión de R_k es $k + 1$, de modo que existen $k + 1$ vectores de la forma $\begin{bmatrix} \mathbf{u} \\ \mathbf{u} \end{bmatrix}$ y uno más, $\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}$. Se tiene que la dimensión de R_{k+1} es $k + 2$ y por tanto R_{k+1} es un código lineal $(2^{k+1}, k + 2)$.

Para la aseveración final, note que los vectores en R_{k+1} se obtienen como combinaciones lineales de los vectores base y por tanto son de la forma

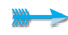
$$\mathbf{v} = c_1\begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_1 \end{bmatrix} + \cdots + c_{k+1}\begin{bmatrix} \mathbf{u}_{k+1} \\ \mathbf{u}_{k+1} \end{bmatrix} + c_{k+2}\begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix}$$

donde $\{\mathbf{u}_1, \ldots, \mathbf{u}_{k+1}\}$ es una base para R_k, **0** y **1** están en $\mathbb{Z}_2^{2^k}$, y cada c_i es 0 o 1. Suponga $\mathbf{v} \neq \mathbf{0}, \mathbf{1}$ y sea $\mathbf{u} = c_1\mathbf{u}_1 + \cdots + c_{k+1}\mathbf{u}_{k+1}$. (En consecuencia, **u** está en R_k.) Si $c_{k+2} = 0$, entonces $\mathbf{u} \neq \mathbf{0}, \mathbf{1}$, de modo que, por la hipótesis de inducción, **u** tiene peso 2^{k-1}. Pero entonces **v** tiene peso $2 \cdot 2^{k-1} = 2^k$. Si $c_{k+2} = 1$, entonces **v** tiene la forma

$$\mathbf{v} = \begin{bmatrix} \mathbf{u} \\ \mathbf{u} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{u} \\ \mathbf{u} + \mathbf{1} \end{bmatrix} = \begin{bmatrix} \mathbf{u} \\ \overline{\mathbf{u}} \end{bmatrix}$$

donde **u** está en R_k. Puesto que

$$w(\overline{\mathbf{u}}) = 2^k - w(\mathbf{u})$$


(¿por qué?), se tiene

$$w(\mathbf{v}) = w(\mathbf{u}) + w(\overline{\mathbf{u}}) = 2^k$$

como se requiere. Esto completa la inducción y se concluye que el teorema es verdadero para toda $n \geq 1$.

Como se mencionó, el *Mariner 9* requirió un código con $64 = 2^6$ vectores. Por el Teorema 6.35, el código de Reed-Muller R_5 tiene dimensión 6 sobre $\mathbb{Z}_2$. Como verá en el siguiente capítulo, también es capaz de detectar y corregir múltiples errores. Es por esto que la NASA usó el código de Reed-Muller para transmitir las fotografías del *Mariner*. Los ejercicios 35-38 exploran más aspectos de esta importante clase de códigos.

Ejercicios 6.7

Ecuaciones diferenciales lineales homogéneas

En los ejercicios 1-12, encuentre la solución de la ecuación diferencial que satisfaga la condición de frontera dada.

1. $y' - 3y = 0, y(1) = 2$

2. $x' + x = 0, x(1) = 1$

3. $y'' - 7y' + 12y = 0, y(0) = y(1) = 1$

4. $x'' + x' - 12x = 0, x(0) = 0, x'(0) = 1$

5. $f'' - f' - f = 0, f(0) = 0, f(1) = 1$

6. $g'' - 2g = 0, g(0) = 1, g(1) = 0$

7. $y'' - 2y' + y = 0, y(0) = y(1) = 1$

8. $x'' + 4x' + 4x = 0, x(0) = 1, x'(0) = 1$

9. $y'' - k^2y = 0, k \neq 0, y(0) = y'(0) = 1$

10. $y'' - 2ky' + k^2y = 0, k \neq 0, y(0) = 1, y(1) = 0$

11. $f'' - 2f' + 5f = 0, f(0) = 1, f(\pi/4) = 0$

12. $h'' - 4h' + 5h = 0, h(0) = 0, h'(0) = -1$

13. Una cepa de bacterias tiene una tasa de crecimiento que es proporcional al tamaño de la población. Inicialmente, existen 100 bacterias; después de 3 horas, existen 1600.

(a) Si $p(t)$ denota el número de bacterias después de t horas, encuentre una fórmula para $p(t)$.

(b) ¿Cuánto tarda la población en duplicarse?

(c) ¿Cuándo la población llegará a un millón?

CAS **14.** La tabla 6.2 muestra la población de Estados Unidos a intervalos de 10 años para los años 1900-2000.

(a) Si supone un modelo de crecimiento exponencial, use los datos para 1900 y 1910 para encontrar una fórmula para $p(t)$, la población en el año t. [*Sugerencia:* sea $t = 0$ en 1900 y $t = 1$ en 1910.] ¿Con cuánta precisión su fórmula calcula la población estadounidense en 2000?

(b) Repita el inciso (a), pero use los datos para los años 1970 y 1980 para determinar $p(t)$. ¿Este planteamiento ofrece una mejor aproximación para el año 2000?

(c) ¿Qué puede concluir acerca del crecimiento poblacional estadounidense?

Tabla 6.2

Año	Población (en millones)
1900	76
1910	92
1920	106
1930	123
1940	131
1950	150
1960	179
1970	203
1980	227
1990	250
2000	281

Fuente: U.S. Bureau of the Census

15. La vida media del radio 226 es de 1590 años. Suponga que comienza con una muestra de radio 226 cuya masa es de 50 mg.

(a) Encuentre una fórmula para la masa $m(t)$ que permanece después de t años y use esta fórmula para predecir la masa restante después de 1000 años.

(b) ¿Cuándo quedarán sólo 10 mg?

16. La *datación con radiocarbono* es un método que usan los científicos para estimar la edad de objetos antiguos que alguna vez fueron materia viva, como hueso, piel, madera o papel. Todos ellos contienen carbono, una proporción del cual es carbono 14, un isótopo radiactivo que se forma continuamente en la atmósfera superior. Puesto que los organismos vivos asimilan carbono radiactivo junto con otros átomos de carbono, la proporción entre las dos formas permanece constante. Sin embargo, cuando un organismo muere, el carbono 14 en sus células decae y ya no se reemplaza. El carbono 14 tiene una vida media conocida de 5730 años, de modo que, al medir la concentración de carbono 14 en un objeto, los científicos pueden determinar su edad aproximada.

Una de las aplicaciones más exitosas de la datación con radiocarbono fue la determinación de la edad del monumento Stonehenge en Inglaterra (figura 6.25). Se descubrió que las muestras tomadas de los restos de los postes de madera tenían una concentración de carbono 14 que era 45% la que se encuentra en el material vivo. ¿Cuál es la edad estimada de dichos postes?



Figura 6.25
Stonehenge

17. Una masa se une a un resorte, como en el ejemplo 6.92. En el tiempo $t = 0$ segundos, el resorte se estira hasta una longitud de 10 cm por abajo de su posición en reposo. El resorte se libera y se observa que su longitud 10 segundos después es de 5 cm. Encuentre una fórmula para la longitud del resorte en el tiempo t segundos.

18. Una masa de 50 g se une a un resorte, como en el ejemplo 6.92. Si el periodo de oscilación es 10 segundos, encuentre la constante del resorte.

19. Un péndulo consiste de una masa, llamada *lenteja*, que está fija en el extremo de un resorte de longitud L (vea la figura 6.26). Cuando la lenteja se aleja de su posición en reposo y se libera, se balancea de ida y vuelta. El tiempo que tarda el péndulo en ir desde su posición derecha más lejana hasta su posición izquierda más lejana y de vuelta a su siguiente posición derecha más lejana se llama *periodo* del péndulo.

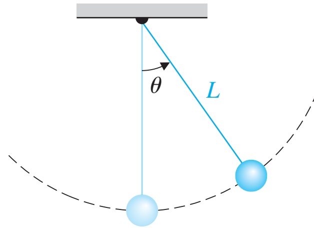


Figura 6.26

Sea $\theta = \theta(t)$ el ángulo del péndulo desde la vertical. Se puede demostrar que si no hay resistencia entonces, cuando θ es pequeño, se satisface la ecuación diferencial

$$\theta'' + \frac{g}{L}\theta = 0$$

donde g es la constante de aceleración debida a la gravedad, aproximadamente 9.7 m/s^2. Suponga que $L = 1$ m y que el péndulo está en reposo (es decir, $\theta = 0$) en el tiempo $t = 0$ segundos. Entonces la lenteja se lleva a la derecha a un ángulo de θ_1 radianes y se libera.

(a) Encuentre el periodo del péndulo.

(b) ¿El periodo depende del ángulo θ_1 desde el que se libera el péndulo? Esta pregunta la planteó y respondió Galileo en 1638. [Galileo Galilei (1564-1642) estudió medicina en la Universidad de Pisa, pero su verdadero interés siempre fue la matemática. En 1592 se designó a Galileo como profesor de matemáticas en la Universidad de Padua, en Venecia, donde impartió principalmente geometría y astronomía. Fue el primero en usar un telescopio para observar las estrellas y los planetas y, al hacerlo, produjo datos experimentales que apoyaban la visión copernicana de que los planetas daban vuelta alrededor del Sol y no de la Tierra. Por esto, Galileo compareció ante la Inquisición, se le puso en arresto domiciliario y se le prohibió publicar sus resultados. Mientras estaba en arresto domiciliario pudo escribir su investigación acerca de los objetos que caen y los péndulos. Sus notas se sacaron de contrabando de Italia y se publicaron en 1638 como *Discursos sobre dos nuevas ciencias*.]

20. Demuestre que el conjunto solución S de la ecuación diferencial de segundo orden $y'' + ay' + by = 0$ es un subespacio de $\mathscr{F}$.

21. Demuestre el Teorema 6.33(b).

22. Demuestre que $e^{pt}\cos qt$ y $e^{pt}\operatorname{sen} qt$ son linealmente independientes.

Códigos lineales

¿Cuál de los códigos de los ejercicios 23-30 son códigos lineales?

23. $C = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

24. $C = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$

25. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

26. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

27. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

28. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

29. El código de paridad par E_n (vea el ejercicio 18 de la sección 5.5.)

30. El código de paridad impar O_n que consiste de todos los vectores en $\mathbb{Z}_2^n$ con peso impar.

31. Encuentre las otras dos bases para el código C_1 en el ejemplo 6.94.

32. **(a)** Si un código lineal (9, 4) tiene matriz generadora G y matriz de control de paridad P, ¿cuáles son las dimensiones de G y P?

(b) Repita el inciso (a) para un código lineal (n, k).

33. Para un código lineal C, demuestre que $(C^\perp)^\perp = C$ sin usar matrices.

34. Si C es un código lineal (n, k) que es autodual, demuestre que n debe ser par. [*Sugerencia:* use el análogo en $\mathbb{Z}_2^n$ del Teorema 5.13.]

35. Escriba los vectores en el código R_3 de Reed-Muller.

36. Defina inductivamente una familia de matrices del modo siguiente: $G_0 = [1]$ y para $n \geq 1$,

$$G_n = \left[\begin{array}{c|c} G_{n-1} & \mathbf{0} \\ \hline G_{n-1} & \mathbf{1} \end{array}\right]$$

donde $\mathbf{0}$ es un vector cero y $\mathbf{1}$ es un vector que consiste por completo de números 1.

(a) Escriba G_1, G_2 y G_3.
(b) Mediante inducción, demuestre que para todo $n \geq 0$, G_n es una matriz generadora para el código R_n de Reed-Muller.

37. Encuentre una matriz de control de paridad para R_2.

38. Encuentre una matriz de control de paridad para R_3.

39. Demuestre que para un código lineal C, todos los vectores código tienen peso par o exactamente la mitad de ellos lo tienen. [*Sugerencia:* sea E el conjunto de vectores en C con peso par y O el conjunto de vectores en C con peso impar. Si O no está vacío, sea $\mathbf{c}_o$ que está en O y considere $O' = \{\mathbf{c}_o + \mathbf{e} : \mathbf{e} \text{ en } E\}$. Demuestre que $O' = O$.]

Repaso del capítulo

Definiciones y conceptos clave

base, 464
base estándar, 465
combinación lineal de vectores, 451
composición de transformaciones lineales, 495
dimensión, 471
espacio vectorial, 447
generador de un conjunto de vectores, 456
inyectiva, 506
isomorfismo, 511
kernel de una transformación lineal, 500
matriz de cambio de base, 483
matriz de una transformación lineal, 516
nulidad de una transformación lineal, 502
rango de una transformación lineal, 500
rank de una transformación lineal, 502
sobreyectiva, 506
subespacio, 452
subespacio cero, 455
subespacio trivial, 455
teorema de la base, 471
teorema del rank, 504
teorema fundamental de las matrices invertibles, 530
transformación cero, 492
transformación identidad, 492
transformación lineal, 490
transformación lineal diagonalizable, 527
transformación lineal invertible, 496
vector, 447
vector coordenado, 467
vectores linealmente dependientes, 461, 464
vectores linealmente independientes, 461, 464

Preguntas de repaso

1. Marque cada uno de los siguientes enunciados como verdadero o falso.
(a) Si $V = \text{gen}(\mathbf{v}_1, \ldots, \mathbf{v}_n)$, entonces todo conjunto generador para V contiene al menos n vectores.
(b) Si $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ es un conjunto de vectores linealmente independientes, entonces también lo es $\{\mathbf{u} + \mathbf{v}, \mathbf{v} + \mathbf{w}, \mathbf{u} + \mathbf{w}\}$.
(c) M_{22} tiene una base que consiste de matrices invertibles.
(d) M_{22} tiene una base que consiste de matrices cuya traza es cero.
(e) La transformación $T : \mathbb{R}^n \to \mathbb{R}$ definida por $T(\mathbf{x}) = \|\mathbf{x}\|$ es una transformación lineal.
(f) Si $T : V \to W$ es una transformación lineal y $\dim V \neq \dim W$, entonces T no puede ser tanto inyectiva como sobreyectiva.
(g) Si $T : V \to W$ es una transformación lineal y $\ker(T) = V$, entonces $W = \{\mathbf{0}\}$.
(h) Si $T : M_{33} \to \mathscr{P}_4$ es una transformación lineal y nulidad$(T) = 4$, entonces T es sobreyectiva.
(i) El espacio vectorial $V = \{p(x) \text{ en } \mathscr{P}_4 : p(1) = 0\}$ es isomórfico a $\mathscr{P}_3$.
(j) Si $I : V \to V$ es la transformación identidad, entonces la matriz $[I]_{\mathcal{C} \leftarrow \mathcal{B}}$ es la matriz identidad para cualesquiera bases $\mathcal{B}$ y $\mathcal{C}$ de V.

En las preguntas 2-5, determine si W es un subespacio de V.

2. $V = \mathbb{R}^2$, $W = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x^2 + 3y^2 = 0 \right\}$

3. $V = M_{22}$, $W = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a + b = c + d = a + c = b + d \right\}$

4. $V = \mathscr{P}_3$, $W = \{p(x) \text{ en } \mathscr{P}_3 : x^3 p(1/x) = p(x)\}$

5. $V = \mathscr{F}$, $W = \{f \text{ en } \mathscr{F} : f(x + \pi) = f(x) \text{ para toda } x\}$

6. Determine si $\{1, \cos 2x, 3\text{sen}^2 x\}$ es linealmente dependiente o independiente.

7. Sean A y B matrices de $n \times n$ distintas de cero tales que A es simétrica y B es antisimétrica. Pruebe que $\{A, B\}$ es linealmente independiente.

En las preguntas 8 y 9, encuentre una base para W y enuncie la dimensión de W.

8. $W = \left\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} : a + d = b + c\right\}$

9. $W = \{p(x) \text{ en } \mathscr{P}_5 : p(-x) = p(x)\}$

10. Encuentre las matrices de cambio de base $P_{\mathcal{C}\leftarrow\mathcal{B}}$ y $P_{\mathcal{B}\leftarrow\mathcal{C}}$ con respecto a las bases $\mathcal{B} = \{1, 1 + x, 1 + x + x^2\}$ y $\mathcal{C} = \{1 + x, x + x^2, 1 + x^2\}$ de $\mathscr{P}_2$.

En las preguntas 11-13, determine si T es una transformación lineal.

11. $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida mediante $T(\mathbf{x}) = \mathbf{y}\mathbf{x}^T\mathbf{y}$, donde $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

12. $T : M_{nn} \to M_{nn}$ definida mediante $T(A) = A^T A$

13. $T : \mathscr{P}_n \to \mathscr{P}_n$ definida mediante $T(p(x)) = p(2x - 1)$

14. Si $T : \mathscr{P}_2 \to M_{22}$ es una transformación lineal tal que $T(1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $T(1 + x) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ y $T(1 + x + x^2) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$, encuentre $T(5 - 3x + 2x^2)$.

15. Encuentre la nulidad de la transformación lineal $T : M_{nn} \to \mathbb{R}$ definida mediante $T(A) = \text{tr}(A)$.

16. Sea W el espacio vectorial de las matrices triangulares superiores de 2×2.

(**a**) Encuentre una transformación lineal $T : M_{22} \to M_{22}$ tal que $\ker(T) = W$.

(**b**) Encuentre una transformación lineal $T : M_{22} \to M_{22}$ tal que $\text{rango}(T) = W$.

17. Encuentre la matriz $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ de la transformación lineal T en la pregunta 14 con respecto a las bases estándar $\mathcal{B} = \{1, x, x^2\}$ de $\mathscr{P}_2$ y $\mathcal{C} = \{E_{11}, E_{12}, E_{21}, E_{22}\}$ de M_{22}.

18. Sea $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ un conjunto de vectores en un espacio vectorial V con la propiedad de que todo vector en V puede escribirse como una combinación lineal de $\mathbf{v}_1, \ldots, \mathbf{v}_n$ en exactamente una forma. Pruebe que S es una base para V.

19. Si $T : U \to V$ y $S : V \to W$ son transformaciones lineales tales que $\text{rango}(T) \subseteq \ker(S)$, ¿qué puede deducir acerca de $S \circ T$?

20. Sea $T : V \to V$ una transformación lineal, y sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base para V tal que $\{T(\mathbf{v}_1), \ldots, T(\mathbf{v}_n)\}$ también es una base para V. Pruebe que T es invertible.

7 Distancia y aproximación

Una línea recta puede ser la distancia más corta entre dos puntos, pero de ninguna forma es la más interesante.

–Doctor Who en "The Time Monster" de Robert Sloman
BBC, 1972

Aunque esto pueda parecer paradójico, toda ciencia exacta está dominada por la idea de la aproximación.

–Bertrand Russell
en W. H. Auden y L. Kronnenberger, eds.,
The Viking Book of Aphorisms,
Viking, 1962, p. 263

7.0 Introducción: geometría de taxi

Se vive en un mundo euclidiano tridimensional y, por tanto, los conceptos de la geometría euclidiana gobiernan la forma en que se ve el mundo. En particular, imagine que detiene a la gente en la calle y que les pide llenar el espacio en blanco de la siguiente oración: "la distancia más corta entre dos puntos es ______". Con mucha seguridad responderán "una línea recta". Sin embargo, existen otras nociones de distancia igualmente sensibles e intuitivas. Al permitirse pensar en "distancia" en una forma más flexible, abrirá la puerta a la posibilidad de tener una "distancia" entre polinomios, funciones, matrices y muchos otros objetos que surgen en el álgebra lineal.

En esta sección descubrirá un tipo de "distancia" que es tan real como la distancia en línea recta a que está acostumbrado en la geometría euclidiana (la que es consecuencia del teorema de Pitágoras). Como verá, este nuevo tipo de "distancia" todavía se comporta en algunas formas familiares.

Suponga que está de pie en una intersección en una ciudad y trata de llegar a un restaurante en otra intersección. Si le pregunta a alguien qué tan lejos está el restaurante, es improbable que esa persona mida la distancia "a vuelo de pájaro" (esto es: usando la versión euclidiana de distancia). En vez de ello, la respuesta será algo como "está a cinco cuadras". Dado que esta es la forma como los conductores de taxi miden la distancia, a esta noción de "distancia" se le referirá como ***distancia de taxi***.

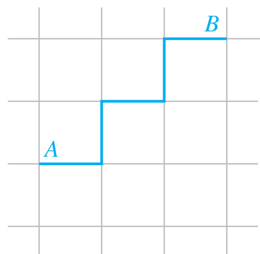


Figura 7.1
Distancia de taxi

La figura 7.1 muestra un ejemplo de distancia de taxi. La ruta más corta de A a B requiere recorrer los lados de cinco cuadras de la ciudad. Note que, aunque hay más de una ruta de A a B, todas las rutas más cortas requieren tres movimientos horizontales y dos verticales, donde un "movimiento" corresponde al lado de una cuadra de la ciudad. (¿Cuántas rutas más cortas hay de A a B?) Por tanto, la distancia de taxi de A a B es 5.

Al idealizar esta situación, se supondrá que todas las cuadras son cuadrados unitarios, y se usará la notación $d_t(A, B)$ para la distancia de taxi de A a B.

Problema 1 Encuentre la distancia de taxi entre los siguientes pares de puntos:

(a) $(1, 2)$ y $(5, 5)$ (b) $(2, 4)$ y $(3, -2)$
(c) $(0, 0)$ y $(-4, -3)$ (d) $(-2, 3)$ y $(1, 3)$
(e) $(1, \frac{1}{2})$ y $(-\frac{3}{2}, \frac{3}{2})$ (f) $(2.5, 4.6)$ y $(3.1, 1.5)$

Problema 2 ¿Cuál de las siguientes es la fórmula correcta para la distancia de taxi $d_t(A, B)$ entre $A = (a_1, a_2)$ y $B = (b_1, b_2)$?

(a) $d_t(A, B) = (a_1 - b_1) + (a_2 - b_2)$
(b) $d_t(A, B) = (|a_1| - |b_1|) + (|a_2| - |b_2|)$
(c) $d_t(A, B) = |a_1 - b_1| + |a_2 - b_2|$

Es posible definir la ***norma de taxi*** de un vector **v** como

$$\|\mathbf{v}\|_t = d_t(\mathbf{v}, \mathbf{0})$$

Problema 3 Encuentre $\|\mathbf{v}\|_t$ para los siguientes vectores:

(a) $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ (b) $\mathbf{v} = \begin{bmatrix} 6 \\ -4 \end{bmatrix}$

(c) $\mathbf{v} = \begin{bmatrix} -3 \\ -6 \end{bmatrix}$ (d) $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

Problema 4 Demuestre que el Teorema 1.3 es verdadero para la norma de taxi.

Problema 5 Verifique la desigualdad del triángulo (Teorema 1.5) usando la norma de taxi y los siguientes pares de vectores:

(a) $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ (b) $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$

Problema 6 Demuestre que la desigualdad del triángulo es verdadera, en general, para la norma de taxi.

En geometría euclidiana puede definirse una circunferencia de radio r, con centro en el origen, como el conjunto de todos los **x** tales que $\|\mathbf{x}\| = r$. Por analogía, puede definir una ***circunferencia de taxi*** de radio r, con centro en el origen, como el conjunto de todos los **x** tales que $\|\mathbf{x}\|_t = r$.

Problema 7 Dibuje circunferencias de taxi con centro en el origen y los siguientes radios:

(a) $r = 3$ (b) $r = 4$ (c) $r = 1$

Problema 8 En geometría euclidiana, el valor de π es la mitad de la circunferencia de un círculo unitario (un círculo de radio 1). Defina la ***pi de taxi*** como el número π_t que es la mitad de la circunferencia de un círculo de taxi unitario. ¿Cuál es el valor de π_t?

En geometría euclidiana, el bisector perpendicular de un segmento de recta $\overline{AB}$ puede definirse como el conjunto de todos los puntos que son equidistantes a A y B. Si usa la distancia de taxi en lugar de la distancia euclidiana, es razonable preguntar cómo se ve ahora el bisector perpendicular de un segmento de recta. Para ser precisos: el ***bisector perpendicular de taxi*** de $\overline{AB}$ es el conjunto de todos los puntos X tales que

$$d_t(X, A) = d_t(X, B)$$

Problema 9 Dibuje el bisector perpendicular de taxi de $\overline{AB}$ para los siguientes pares de puntos:

(a) $A = (2, 1), B = (4, 1)$ (b) $A = (-1, 3), B = (-1, -2)$
(c) $A = (1, 1), B = (5, 3)$ (d) $A = (1, 1), B = (5, 5)$

Como ilustra cada uno de estos problemas, la geometría de taxi comparte algunas propiedades con la geometría euclidiana, pero también difiere en algunas formas impre-

sionantes. En este capítulo encontrará muchos otros tipos de distancias y normas, cada una de las cuales es útil por cuenta propia. Se intentará descubrir qué tienen en común y usará dichas propiedades comunes en su beneficio. También explorará diversos problemas de aproximación en los que la noción de "distancia" tiene un papel importante.

7.1 Espacios con producto interno

En el capítulo 1 se definió el producto punto $\mathbf{u} \cdot \mathbf{v}$ de los vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{R}^n$, y a lo largo del libro se ha usado repetidamente. En esta sección se usarán las propiedades del producto punto como un medio para definir la noción general de un *producto interno*. En la siguiente sección se demostrará que los productos internos pueden usarse para definir análogos de "longitud" y "distancia" en espacios vectoriales distintos a $\mathbb{R}^n$.

La siguiente definición es el punto de partida; se basa en las propiedades del producto punto demostradas en el Teorema 1.2.

Definición Un ***producto interno*** sobre un espacio vectorial V es una operación que asigna a cada par de vectores $\mathbf{u}$ y $\mathbf{v}$ en V un número real $\langle \mathbf{u}, \mathbf{v} \rangle$ tal que las siguientes propiedades se cumplen para todos los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ en V y todos los escalares c:

1. $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$
2. $\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$
3. $\langle c\mathbf{u}, \mathbf{v} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle$
4. $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ y $\langle \mathbf{u}, \mathbf{u} \rangle = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$

Un espacio vectorial con un producto interno se llama ***espacio con producto interno***.

Comentario Técnicamente, esta definición especifica un espacio con producto interno *real*, pues supone que V es un espacio vectorial real y que el producto interno de dos vectores es un número real. También existen espacios con producto interno *complejos*, pero su definición es un tanto diferente. (Vea la Exploración: "Vectores y matrices con entradas complejas" al final de esta sección.)

Ejemplo 7.1

$\mathbb{R}^n$ es un espacio con producto interno con $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$. Las propiedades (1) a (4) se verificaron como el Teorema 1.2.

El producto punto no es el único producto interno que puede definirse en $\mathbb{R}^n$.

Ejemplo 7.2

Sean $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ dos vectores en $\mathbb{R}^2$. Demuestre que

$$\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1v_1 + 3u_2v_2$$

define un producto interno.

Solución Se deben verificar las propiedades (1) a (4). La propiedad (1) se cumple porque

$$\langle \mathbf{u}, \mathbf{v} \rangle = 2u_1v_1 + 3u_2v_2 = 2v_1u_1 + 3v_2u_2 = \langle \mathbf{v}, \mathbf{u} \rangle$$

A continuación, sea $\mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$. Se comprueba que

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle &= 2u_1(v_1 + w_1) + 3u_2(v_2 + w_2) \\ &= 2u_1v_1 + 2u_1w_1 + 3u_2v_2 + 3u_2w_2 \\ &= (2u_1v_1 + 3u_2v_2) + (2u_1w_1 + 3u_2w_2) \\ &= \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle \end{aligned}$$

lo que prueba la propiedad (2).

Si c es un escalar, entonces

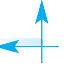
$$\begin{aligned} \langle c\mathbf{u}, \mathbf{v} \rangle &= 2(cu_1)v_1 + 3(cu_2)v_2 \\ &= c(2u_1v_1 + 3u_2v_2) \\ &= c\langle \mathbf{u}, \mathbf{v} \rangle \end{aligned}$$

lo que verifica la propiedad (3).

Finalmente,

$$\langle \mathbf{u}, \mathbf{u} \rangle = 2u_1u_1 + 3u_2u_2 = 2u_1^2 + 3u_2^2 \geq 0$$

y es claro que $\langle \mathbf{u}, \mathbf{u} \rangle = 2u_1^2 + 3u_2^2 = 0$ si y sólo si $u_1 = u_2 = 0$ (esto es, si y sólo si $\mathbf{u} = \mathbf{0}$). Esto verifica la propiedad (4) y completa la demostración de que $\langle \mathbf{u}, \mathbf{v} \rangle$, como se definió, es un producto interno.



El ejemplo 7.2 puede generalizarse para demostrar que si $w_1, \ldots, w_n$ son escalares *positivos* y

$$\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

son vectores en $\mathbb{R}^n$, entonces

$$\langle \mathbf{u}, \mathbf{v} \rangle = w_1u_1v_1 + \cdots + w_nu_nv_n \tag{1}$$

define un producto interno sobre $\mathbb{R}^n$, llamado ***producto punto ponderado***. Si alguno de los pesos w_i es negativo o cero, entonces la ecuación (1) no define un producto interno. (Vea los ejercicios 13 y 14.)

Recuerde que el producto punto puede expresarse como $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v}$. Observe que es posible escribir el producto punto ponderado en la ecuación (1) como

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T W \mathbf{v}$$

donde W es la matriz diagonal de $n \times n$

$$W = \begin{bmatrix} w_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & w_n \end{bmatrix}$$

El siguiente ejemplo generaliza aún más este tipo de producto interno.

Ejemplo 7.3

Sea A una matriz simétrica definida positiva de $n \times n$ (vea la sección 5.5) y sean $\mathbf{u}$ y $\mathbf{v}$ vectores en $\mathbb{R}^n$. Demuestre que

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$$

define un producto interno.

Solución Compruebe que

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= \mathbf{u}^T A \mathbf{v} = \mathbf{u} \cdot A\mathbf{v} = A\mathbf{v} \cdot \mathbf{u} \\ &= A^T \mathbf{v} \cdot \mathbf{u} = (\mathbf{v}^T A)^T \cdot \mathbf{u} = \mathbf{v}^T A \mathbf{u} = \langle \mathbf{v}, \mathbf{u} \rangle \end{aligned}$$

Además,

$$\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \mathbf{u}^T A(\mathbf{v} + \mathbf{w}) = \mathbf{u}^T A \mathbf{v} + \mathbf{u}^T A \mathbf{w} = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$$

y

$$\langle c\mathbf{u}, \mathbf{v} \rangle = (c\mathbf{u})^T A \mathbf{v} = c(\mathbf{u}^T A \mathbf{v}) = c\langle \mathbf{u}, \mathbf{v} \rangle$$

Finalmente, puesto que A es definida positiva, $\langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u}^T A \mathbf{u} > 0$ para toda $\mathbf{u} \neq \mathbf{0}$, de modo que $\langle \mathbf{u}, \mathbf{u} \rangle = \mathbf{u}^T A \mathbf{u} = \mathbf{0}$, si y sólo si $\mathbf{u} = \mathbf{0}$. Esto establece la última propiedad.

Para ilustrar el ejemplo 7.3, sea $A = \begin{bmatrix} 4 & -2 \\ -2 & 7 \end{bmatrix}$. Entonces

$$\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v} = [u_1 \ \ u_2] \begin{bmatrix} 4 & -2 \\ -2 & 7 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = 4u_1v_1 - 2u_1v_2 - 2u_2v_1 + 7u_2v_2$$

La matriz A es definida positiva, por el Teorema 5.24, pues sus eigenvalores son 3 y 8. Por tanto, $\langle \mathbf{u}, \mathbf{v} \rangle$ define un producto interno sobre $\mathbb{R}^2$.

Ahora se definirán algunos productos internos sobre espacios vectoriales distintos a $\mathbb{R}^n$.

Ejemplo 7.4

En $\mathscr{P}_2$, sean $p(x) = a_0 + a_1x + a_2x^2$ y $q(x) = b_0 + b_1x + b_2x^2$. Demuestre que

$$\langle p(x), q(x) \rangle = a_0b_0 + a_1b_1 + a_2b_2$$

define un producto interno sobre $\mathscr{P}_2$. (Por ejemplo, si $p(x) = 1 - 5x + 3x^2$ y $q(x) = 6 + 2x - x^2$, entonces $\langle p(x), q(x) \rangle = 1 \cdot 6 + (-5) \cdot 2 + 3 \cdot (-1) = -7$.)

Solución Puesto que $\mathscr{P}_2$ es isomórfico a $\mathbb{R}^3$, sólo es necesario demostrar que el producto punto en $\mathbb{R}^3$ es un producto interno, lo que ya se estableció.



Ejemplo 7.5

Sean f y g que están en $\mathscr{C}[a, b]$, el espacio vectorial de todas las funciones continuas sobre el intervalo cerrado $[a, b]$. Demuestre que

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx$$

define un producto interno sobre $\mathscr{C}[a, b]$.

Solución Se tiene

$$\langle f, g \rangle = \int_a^b f(x)g(x)\, dx = \int_a^b g(x)f(x)\, dx = \langle g, f \rangle$$

Además, si h está en $\mathscr{C}[a, b]$, entonces

$$\begin{aligned}\langle f, g + h \rangle &= \int_a^b f(x)(g(x) + h(x))\, dx \\ &= \int_a^b (f(x)g(x) + f(x)h(x))\, dx \\ &= \int_a^b f(x)g(x)\, dx + \int_a^b f(x)h(x)\, dx \\ &= \langle f, g \rangle + \langle f, h \rangle\end{aligned}$$

Si c es un escalar, entonces

$$\begin{aligned}\langle cf, g \rangle &= \int_a^b cf(x)g(x)\, dx \\ &= c\int_a^b f(x)g(x)\, dx \\ &= c\langle f, g \rangle\end{aligned}$$

Finalmente, $\langle f, f \rangle = \int_a^b (f(x))^2\, dx \geq 0$, y de un teorema del cálculo se tiene que, puesto que f es continua, $\langle f, f \rangle = \int_a^b (f(x))^2\, dx = 0$ si y sólo si f es la función cero. Por tanto, $\langle f, g \rangle$ es un producto interno sobre $\mathscr{C}[a, b]$.

El ejemplo 7.5 también define un producto interno sobre cualquier *subespacio* de $\mathscr{C}[a, b]$. Por ejemplo, podría restringir la atención a los polinomios definidos en el intervalo $[a, b]$. Suponga que considera $\mathscr{P}[0, 1]$, el espacio vectorial de todos los polinomios sobre el intervalo $[0, 1]$. Entonces, al usar el producto interno del ejemplo 7.5, se tiene

$$\begin{aligned}\langle x^2, 1 + x \rangle &= \int_0^1 x^2(1 + x)\, dx = \int_0^1 (x^2 + x^3)\, dx \\ &= \left[\frac{x^3}{3} + \frac{x^4}{4}\right]_0^1 = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}\end{aligned}$$

Propiedades de los productos internos

El siguiente teorema resume algunas propiedades adicionales que son consecuencia de la definición de producto interno.

Teorema 7.1 Sean $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ vectores en un espacio con producto interno V y sea c un escalar.

a. $\langle \mathbf{u} + \mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$
b. $\langle \mathbf{u}, c\mathbf{v}\rangle = c\langle \mathbf{u}, \mathbf{v}\rangle$
c. $\langle \mathbf{u}, \mathbf{0}\rangle = \langle \mathbf{0}, \mathbf{v}\rangle = 0$

Demostración Se demuestra la propiedad (a) y las demostraciones de las propiedades (b) y (c) se dejan como los ejercicios 23 y 24. Con referencia a la definición del producto interno, se tiene

$$\begin{aligned} \langle \mathbf{u} + \mathbf{v}, \mathbf{w}\rangle &= \langle \mathbf{w}, \mathbf{u} + \mathbf{v}\rangle && \text{por (1)} \\ &= \langle \mathbf{w}, \mathbf{u}\rangle + \langle \mathbf{w}, \mathbf{v}\rangle && \text{por (2)} \\ &= \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle && \text{por (1)} \end{aligned}$$

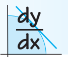
Longitud, distancia y ortogonalidad

En un espacio con producto interno puede definir la longitud de un vector, la distancia entre vectores y los vectores ortogonales, como se hizo en la sección 1.2. Simplemente tiene que sustituir cada uso del producto punto $\mathbf{u}\cdot\mathbf{v}$ por el más general producto interno $\langle \mathbf{u}, \mathbf{v}\rangle$.

Definición Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en un espacio con producto interno V.

1. La ***longitud*** (o ***norma***) de $\mathbf{v}$ es $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v}\rangle}$.
2. La ***distancia*** entre $\mathbf{u}$ y $\mathbf{v}$ es $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$.
3. $\mathbf{u}$ y $\mathbf{v}$ son ***ortogonales*** si $\langle \mathbf{u}, \mathbf{v}\rangle = 0$.

Note que $\|\mathbf{v}\|$ siempre está definido, pues $\langle \mathbf{v}, \mathbf{v}\rangle \geq 0$ por la definición de producto interno, de modo que puede sacar la raíz cuadrada de esta cantidad no negativa. Como en $\mathbb{R}^n$ un vector de longitud 1 se llama ***vector unitario***. La ***esfera unitaria*** en V es el conjunto S de todos los vectores unitarios en V.



Ejemplo 7.6 Considere el producto interno sobre $\mathscr{C}[0, 1]$ dado en el ejemplo 7.5. Si $f(x) = x$ y $g(x) = 3x - 2$, encuentre

(a) $\|f\|$ (b) $d(f, g)$ (c) $\langle f, g\rangle$

Solución (a) Se encuentra que

$$\langle f, f\rangle = \int_0^1 f^2(x)\,dx = \int_0^1 x^2\,dx = \frac{x^3}{3}\bigg]_0^1 = \frac{1}{3}$$

de modo que $\|f\| = \sqrt{\langle f, f\rangle} = 1/\sqrt{3}$.

(b) Dado que $d(f, g) = \|f - g\| = \sqrt{\langle f - g, f - g\rangle}$ y

$$f(x) - g(x) = x - (3x - 2) = 2 - 2x = 2(1 - x)$$

se tiene
$$\langle f - g, f - g\rangle = \int_0^1 (f(x) - g(x))^2\,dx = \int_0^1 4(1 - 2x + x^2)\,dx$$
$$= 4\left[x - x^2 + \frac{x^3}{3}\right]_0^1 = \frac{4}{3}$$

Al combinar estos hechos, se ve que $d(f, g) = \sqrt{4/3} = 2/\sqrt{3}$.

(c) Calcule

$$\langle f, g\rangle = \int_0^1 f(x)g(x)\,dx = \int_0^1 x(3x - 2)\,dx = \int_0^1 (3x^2 - 2x)\,dx = [x^3 - x^2]_0^1 = 0$$

Por tanto, f y g son ortogonales.

Es importante recordar que la "distancia" entre f y g en el ejemplo 7.6 *no* se refiere a alguna medida relacionada con las gráficas de dichas funciones. Tampoco el hecho de que f y g sean ortogonales significa que sus gráficas se intersequen en ángulos rectos. Simplemente se aplica la definición de un producto interno particular. Sin embargo, al hacerlo, uno debe guiarse por las correspondientes nociones en $\mathbb{R}^2$ y $\mathbb{R}^3$, donde el producto interno es el producto punto. La geometría del espacio euclidiano todavía puede guiarlo aquí, aun cuando no pueda visualizar las cosas de la misma forma.

Ejemplo 7.7

Con el producto interno sobre $\mathbb{R}^2$ definido en el ejemplo 7.2, dibuje una gráfica de la esfera (círculo) unitaria.

Solución Si $\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$, entonces $\langle \mathbf{x}, \mathbf{x}\rangle = 2x^2 + 3y^2$. Puesto que la esfera (circunferencia) unitaria consiste de todos los $\mathbf{x}$ tales que $\|\mathbf{x}\| = 1$, se tiene

$$1 = \|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x}\rangle} = \sqrt{2x^2 + 3y^2} \quad \text{o} \quad 2x^2 + 3y^2 = 1$$

Esta es la ecuación de una elipse y su gráfica se muestra en la figura 7.2.

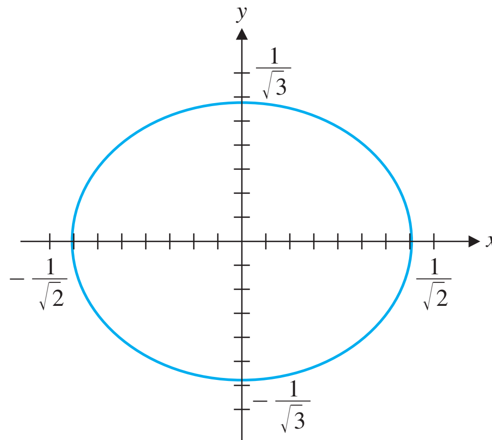


Figura 7.2
Una circunferencia unitaria que es una elipse

En la siguiente sección y en los ejercicios se discutirán las propiedades de longitud, distancia y ortogonalidad. Un resultado que se necesitará en esta sección es la versión generalizada del teorema de Pitágoras, que amplía el Teorema 1.6.

Teorema 7.2 **Teorema de Pitágoras**

Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en un espacio con producto interno V. Entonces $\mathbf{u}$ y $\mathbf{v}$ son ortogonales si y sólo si

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$

Demostración Como se le pedirá demostrar en el ejercicio 32, se tiene

$$\|\mathbf{u} + \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2$$

Se concluye inmediatamente que $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ si y sólo si $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Proyecciones ortogonales y el proceso de Gram-Schmidt

En el capítulo 5 se discutió la ortogonalidad en $\mathbb{R}^n$. La mayor parte de este material se generaliza bastante bien a espacios con producto interno en general. Por ejemplo, un ***conjunto ortogonal*** de vectores en un espacio con producto interno V es un conjunto $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$ de vectores de V tales que $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ siempre que $\mathbf{v}_i \neq \mathbf{v}_j$. Entonces un ***conjunto ortonormal*** de vectores es un conjunto ortogonal de vectores *unitarios*. Una ***base ortogonal*** para un subespacio W de V es justo una base para W que es un conjunto ortogonal; de igual modo, una ***base ortonormal*** para un subespacio W de V es una base para W que es un conjunto ortonormal.

En $\mathbb{R}^n$, el proceso de Gram-Schmidt (Teorema 5.15) muestra que todo subespacio tiene una base ortogonal. Es posible imitar la construcción del proceso de Gram-Schmidt para demostrar que todo subespacio con dimensión finita de un espacio con producto interno tiene una base ortogonal, todo lo que se requiere es sustituir el producto punto por el más general producto interno. Este planteamiento se ilustrará con un ejemplo.  (Compare estos pasos con los dados en el ejemplo 5.13.)

Ejemplo 7.8 Construya una base ortogonal para $\mathscr{P}_2$ con respecto al producto interno

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)\, dx$$

al aplicar el proceso de Gram-Schmidt a la base $\{1, x, x^2\}$.

Solución Sean $\mathbf{x}_1 = 1$, $\mathbf{x}_2 = x$ y $\mathbf{x}_3 = x^2$. Comience por establecer $\mathbf{v}_1 = \mathbf{x}_1 = 1$. A continuación calcule

$$\langle \mathbf{v}_1, \mathbf{v}_1 \rangle = \int_{-1}^{1} dx = x\Big]_{-1}^{1} = 2 \quad \text{y} \quad \langle \mathbf{v}_1, \mathbf{x}_2 \rangle = \int_{-1}^{1} x\, dx = \frac{x^2}{2}\Big]_{-1}^{1} = 0$$

Adrien Marie Legendre (1752-1833) fue un matemático francés que trabajó en astronomía, teoría de números y funciones elípticas. Estuvo involucrado en varias discusiones acaloradas con Gauss. Legendre proporcionó el primer enunciado publicado de la ley de reciprocidad cuadrática en teoría de números en 1765. Sin embargo, Gauss ofreció la primera demostración rigurosa de este resultado en 1801 y reclamó el crédito por el resultado, lo que generó la comprensible irá de Legendre. Luego, en 1806, Legendre ofreció la primera aplicación publicada del método de mínimos cuadrados en un libro acerca de las órbitas de los cometas. Gauss publicó acerca del mismo tema en 1809, pero afirmó que él usaba el método desde 1795, lo que enfureció de nuevo a Legendre.

Por tanto,

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{v}_1, \mathbf{x}_2 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 = x - \frac{0}{2}(1) = x$$

Para encontrar $\mathbf{v}_3$, calcule primero

$$\langle \mathbf{v}_1, \mathbf{x}_3 \rangle = \int_{-1}^{1} x^2\,dx = \frac{x^3}{3}\bigg]_{-1}^{1} = \frac{2}{3}, \qquad \langle \mathbf{v}_2, \mathbf{x}_3 \rangle = \int_{-1}^{1} x^3\,dx = \frac{x^4}{4}\bigg]_{-1}^{1} = 0,$$

$$\langle \mathbf{v}_2, \mathbf{v}_2 \rangle = \int_{-1}^{1} x^2\,dx = \frac{2}{3}$$

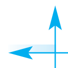
Entonces

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{v}_1, \mathbf{x}_3 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{v}_2, \mathbf{x}_3 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2 = x^2 - \frac{\frac{2}{3}}{2}(1) - \frac{0}{\frac{2}{3}}x = x^2 - \frac{1}{3}$$

Se tiene que $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ es una base ortogonal para $\mathscr{P}_2$ en el intervalo $[-1, 1]$. Los polinomios

$$1, \quad x, \quad x^2 - \tfrac{1}{3}$$

son los primeros tres ***polinomios de Legendre***. Si cada uno de dichos polinomios se dividen por sus longitudes relativas al mismo producto interno, se obtienen ***polinomios normalizados de Legendre*** (vea el ejercicio 41).



Tal como se hizo en la sección 5.2, es posible definir la ***proyección ortogonal*** $\text{proy}_W(\mathbf{v})$ de un vector $\mathbf{v}$ sobre un subespacio W de un espacio con producto interno. Si $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es una base ortogonal para W, entonces

$$\text{proy}_W(\mathbf{v}) = \frac{\langle \mathbf{u}_1, \mathbf{v} \rangle}{\langle \mathbf{u}_1, \mathbf{u}_1 \rangle}\mathbf{u}_1 + \cdots + \frac{\langle \mathbf{u}_k, \mathbf{v} \rangle}{\langle \mathbf{u}_k, \mathbf{u}_k \rangle}\mathbf{u}_k$$

Entonces el ***componente de* v *ortogonal a W*** es el vector

$$\text{perp}_W(\mathbf{v}) = \mathbf{v} - \text{proy}_W(\mathbf{v})$$

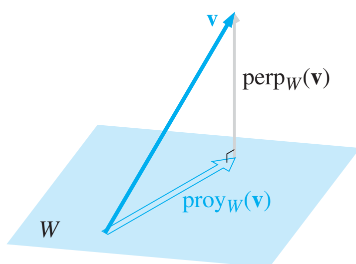


Figura 7.3

Como en el teorema de descomposición ortogonal (Teorema 5.11), $\text{proy}_W(\mathbf{v})$ y $\text{perp}_W(\mathbf{v})$ son ortogonales (vea el ejercicio 43) y, por tanto, esquemáticamente, se tiene la situación que se ilustra en la figura 7.3.

Dichas fórmulas se usarán en las secciones 7.3 y 7.5 cuando se consideren problemas de aproximación, en particular, el problema de cómo aproximar mejor una función dada mediante funciones "amigables". En consecuencia, cualquier ejemplo se diferirá

hasta entonces, cuando tendrá más sentido. El uso inmediato de la proyección ortogonal será demostrar una desigualdad que se encontró por primera vez en el capítulo 1.

Las desigualdades de Cauchy-Schwarz y del triángulo

Las pruebas de las identidades y desigualdades que involucran el producto punto en $\mathbb{R}^n$ se adaptan fácilmente para ofrecer resultados correspondientes en los espacios generales con producto interno. Algunas de ellas se proporcionan en los ejercicios 31-36. En la sección 1.2 se enunció sin demostración la desigualdad de Cauchy-Schwarz, que es importante en muchas ramas de la matemática. Ahora se dará una demostración de este resultado.

Teorema 7.3 **La desigualdad de Cauchy-Schwarz**

Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en un espacio con producto interno V. Entonces

$$|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$$

donde la igualdad se cumple si y sólo si $\mathbf{u}$ y $\mathbf{v}$ son múltiplos escalares uno de otro.

Demostración Si $\mathbf{u} = \mathbf{0}$, entonces la desigualdad en realidad es una igualdad, pues

$$|\langle \mathbf{0}, \mathbf{v}\rangle| = 0 = \|\mathbf{0}\|\,\|\mathbf{v}\|$$

Esta desigualdad la descubrieron varios matemáticos en contextos diferentes. No es de sorprender que el nombre del prolífico Cauchy se vincule con ella. El segundo nombre asociado con este resultado es el de Karl Herman Amandus Schwarz (1843-1921), un matemático alemán que impartió clases en la Universidad de Berlín. Su versión de la desigualdad que lleva su nombre se publicó en 1885 en un ensayo que usó ecuaciones integrales para estudiar superficies de área mínima. Un tercer nombre que también se asocia con este importante resultado es el del matemático ruso Viktor Yakovlevitch Bunyakovsky (1804-1889). Bunyakovsky publicó la desigualdad en 1859, un cuarto de siglo antes que el trabajo de Schwarz acerca del mismo tema. Por tanto, es más adecuado referirse al resultado como la desigualdad de Cauchy-Bunyakovsky-Schwarz.

Si $\mathbf{u} \neq \mathbf{0}$, entonces sea W el subespacio de V generado por $\mathbf{u}$. Puesto que $\text{proy}_W(\mathbf{v}) = \dfrac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}$ y $\text{perp}_W\,\mathbf{v} = \mathbf{v} - \text{proy}_W(\mathbf{v})$ son ortogonales, es posible aplicar el Teorema de Pitágoras para obtener

$$\begin{aligned}\|\mathbf{v}\|^2 &= \|\text{proy}_W(\mathbf{v}) + (\mathbf{v} - \text{proy}_W(\mathbf{v}))\|^2 = \|\text{proy}_W(\mathbf{v}) + \text{perp}_W(\mathbf{v})\|^2 \\ &= \|\text{proy}_W(\mathbf{v})\|^2 + \|\text{perp}_W(\mathbf{v})\|^2\end{aligned} \qquad (2)$$

Se tiene que $\|\text{proy}_W(\mathbf{v})\|^2 \leq \|\mathbf{v}\|^2$. Ahora

$$\|\text{proy}_W(\mathbf{v})\|^2 = \left\langle \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}, \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u} \right\rangle = \left(\frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\right)^2 \langle \mathbf{u}, \mathbf{u}\rangle = \frac{\langle \mathbf{u}, \mathbf{v}\rangle^2}{\langle \mathbf{u}, \mathbf{u}\rangle} = \frac{\langle \mathbf{u}, \mathbf{v}\rangle^2}{\|\mathbf{u}\|^2}$$

así que se tiene

$$\frac{\langle \mathbf{u}, \mathbf{v}\rangle^2}{\|\mathbf{u}\|^2} \leq \|\mathbf{v}\|^2 \quad \text{o, de manera equivalente,} \quad \langle \mathbf{u}, \mathbf{v}\rangle^2 \leq \|\mathbf{u}\|^2\|\mathbf{v}\|^2$$

Al sacar raíces cuadradas, se obtiene $|\langle \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}\|\,\|\mathbf{v}\|$.

Claramente, esta última desigualdad es una igualdad si y sólo si $\|\text{poy}_w(\mathbf{v})\|^2 = |\mathbf{v}|^2$. Por la ecuación (2) esto es verdadero si y sólo si $\text{perp}_W(\mathbf{v}) = \mathbf{0}$ o, de manera equivalente,

$$\mathbf{v} = \text{proy}_W(\mathbf{v}) = \frac{\langle \mathbf{u}, \mathbf{v}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u}$$

Si esto es así, entonces **v** es un múltiplo escalar de **u**. Por el contrario, si $\mathbf{v} = c\mathbf{u}$, entonces

$$\text{perp}_W(\mathbf{v}) = \mathbf{v} - \text{proy}_W(\mathbf{v}) = c\mathbf{u} - \frac{\langle \mathbf{u}, c\mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u} = c\mathbf{u} - \frac{c\langle \mathbf{u}, \mathbf{u}\rangle}{\langle \mathbf{u}, \mathbf{u}\rangle}\mathbf{u} = \mathbf{0}$$

así que la igualdad se mantiene en la desigualdad de Cauchy-Schwarz.

Para una demostración alternativa de esta desigualdad, vea el ejercicio 44. En la Exploración: “Desigualdades geométricas y problemas de optimización”, que sigue a esta sección, se investigarán algunas interesantes consecuencias de la desigualdad de Cauchy-Schwarz. Por el momento, se le usará para demostrar una versión generalizada de la desigualdad del triángulo (Teorema 1.5).

Teorema 7.4 **La desigualdad del triángulo**

Sean **u** y **v** vectores en un espacio con producto interno V. Entonces

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

Demostración A partir de la igualdad que se le pedirá demostrar en el ejercicio 32, se tiene

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2|\langle \mathbf{u}, \mathbf{v}\rangle| + \|\mathbf{v}\|^2 \\ &\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 \quad \text{por Cauchy-Schwarz} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

Al sacar raíces cuadradas se obtiene el resultado.

Ejercicios 7.1

En los ejercicios 1 y 2, sean $\mathbf{u} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$.

1. $\langle \mathbf{u}, \mathbf{v}\rangle$ es el producto interno del ejemplo 7.2. Calcule
   (a) $\langle \mathbf{u}, \mathbf{v}\rangle$ **(b)** $\|\mathbf{u}\|$ **(c)** $d(\mathbf{u}, \mathbf{v})$

2. $\langle \mathbf{u}, \mathbf{v}\rangle$ es el producto interno del ejemplo 7.3 con $A = \begin{bmatrix} 4 & -2 \\ -2 & 7 \end{bmatrix}$. Calcule
   (a) $\langle \mathbf{u}, \mathbf{v}\rangle$ **(b)** $\|\mathbf{u}\|$ **(c)** $d(\mathbf{u}, \mathbf{v})$

3. En el ejercicio 1, encuentre un vector distinto de cero ortogonal a **u**.

4. En el ejercicio 2, encuentre un vector distinto de cero ortogonal a **u**.

En los ejercicios 5 y 6, sean $p(x) = 2 - 3x + x^2$ *y* $q(x) = 1 - 3x^2$. *Calcule*
(a) $\langle p(x), q(x)\rangle$ *(b)* $\|p(x)\|$ *(c)* $d(p(x), q(x))$

5. $\langle p(x), q(x)\rangle$ es el producto interno del ejemplo 7.4.

6. $\langle p(x), q(x)\rangle$ es el producto interno del ejemplo 7.5 sobre el espacio vectorial, $\mathscr{P}_2[0, 1]$.

7. En el ejercicio 5, encuentre un vector distinto de cero ortogonal a $p(x)$.

8. En el ejercicio 6, encuentre un vector distinto de cero ortogonal a $p(x)$.

En los ejercicios 9 y 10, sean $f(x) = \operatorname{sen} x$ *y* $g(x) = \operatorname{sen} x + \cos x$ *en el espacio vectorial* $\mathscr{C}[0, 2\pi]$ *con el producto interno definido por el ejemplo 7.5.*

9. Calcule
   (a) $\langle f, g\rangle$ **(b)** $\|f\|$ **(c)** $d(f, g)$

10. Encuentre un vector distinto de cero ortogonal a f.

11. Sean a, b y c números reales distintos. Demuestre que

$$\langle p(x), q(x)\rangle = p(a)q(a) + p(b)q(b) + p(c)q(c)$$

define un producto interno sobre $\mathcal{P}_2$. [*Sugerencia:* necesitará el hecho de que un polinomio de grado n tiene como máximo n ceros. Vea el Apéndice D.]

12. Repita el ejercicio 5 usando el producto interno del ejercicio 11 con $a = 0, b = 1, c = 2$.

En los ejercicios 13-18, determine cuál de los cuatro axiomas de producto interno no se cumple. Proporcione un ejemplo específico en cada caso.

13. Sean $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ en $\mathbb{R}^2$. Defina $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1$.

14. Sean $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ en $\mathbb{R}^2$. Defina $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 - u_2 v_2$.

15. Sean $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ en $\mathbb{R}^2$. Defina $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_2 + u_2 v_1$.

16. En $\mathcal{P}_2$, defina $\langle p(x), q(x) \rangle = p(0)q(0)$.

17. En $\mathcal{P}_2$, defina $\langle p(x), q(x) \rangle = p(1)q(1)$.

18. En M_{22}, defina $\langle A, B \rangle = \det(AB)$.

En los ejercicios 19 y 20, $\langle \mathbf{u}, \mathbf{v} \rangle$ grafique un producto interno sobre $\mathbb{R}^2$, donde $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. Encuentre una matriz simétrica A tal que $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$.

19. $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1 v_1 + u_1 v_2 + u_2 v_1 + 4u_2 v_2$

20. $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + 2u_1 v_2 + 2u_2 v_1 + 5u_2 v_2$

En los ejercicios 21 y 22, bosqueje el círculo unitario en $\mathbb{R}^2$ para el producto interno dado, donde $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$.

21. $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + \frac{1}{4} u_2 v_2$

22. $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1 v_1 + u_1 v_2 + u_2 v_1 + 4u_2 v_2$

23. Demuestre el Teorema 7.1(b).

24. Demuestre el Teorema 7.1(c).

En los ejercicios 25-29, suponga que **u**, **v** *y* **w** *son vectores en un espacio con producto interno tal que*

$$\langle \mathbf{u}, \mathbf{v} \rangle = 1, \quad \langle \mathbf{u}, \mathbf{w} \rangle = 5, \quad \langle \mathbf{v}, \mathbf{w} \rangle = 0$$
$$\|\mathbf{u}\| = 1, \quad \|\mathbf{v}\| = \sqrt{3}, \quad \|\mathbf{w}\| = 2$$

Evalúe las expresiones en los ejercicios 25-28.

25. $\langle \mathbf{u} + \mathbf{w}, \mathbf{v} - \mathbf{w} \rangle$

26. $\langle 2\mathbf{v} - \mathbf{w}, 3\mathbf{u} + 2\mathbf{w} \rangle$

27. $\|\mathbf{u} + \mathbf{v}\|$

28. $\|2\mathbf{u} - 3\mathbf{v} + \mathbf{w}\|$

29. Demuestre que $\mathbf{u} + \mathbf{v} = \mathbf{w}$. [*Sugerencia:* ¿cómo puede usar las propiedades del producto interno para verificar que $\mathbf{u} + \mathbf{v} - \mathbf{w} = \mathbf{0}$?]

30. Demuestre que en un espacio con producto interno, no puede haber vectores unitarios **u** y **v** con $\langle \mathbf{u}, \mathbf{v} \rangle < -1$.

En los ejercicios 31-36, $\langle \mathbf{u}, \mathbf{v} \rangle$ es un producto interno. En los ejercicios 31-34, demuestre que el enunciado dado es una identidad.

31. $\langle \mathbf{u} + \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2$

32. $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle + \|\mathbf{v}\|^2$

33. $\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = \frac{1}{2}\|\mathbf{u} + \mathbf{v}\|^2 + \frac{1}{2}\|\mathbf{u} - \mathbf{v}\|^2$

34. $\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\|\mathbf{u} + \mathbf{v}\|^2 - \frac{1}{4}\|\mathbf{u} - \mathbf{v}\|^2$

35. Demuestre que $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\|$ si y sólo si **u** y **v** son ortogonales.

36. Demuestre que $d(\mathbf{u}, \mathbf{v}) = \sqrt{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2}$ si y sólo si **u** y **v** son ortogonales.

En los ejercicios 37-40, aplique el proceso de Gram-Schmidt a la base $\mathcal{B}$ para obtener una base ortogonal para el espacio con producto interno V relativo al producto interno dado.

37. $V = \mathbb{R}^2, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$, con el producto interno del ejemplo 7.2

38. $V = \mathbb{R}^2, \mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$, con el producto interno que sigue al ejemplo 7.3

39. $V = \mathcal{P}_2, \mathcal{B} = \{1, 1 + x, 1 + x + x^2\}$, con el producto interno del ejemplo 7.4

40. $V = \mathcal{P}_2[0, 1], \mathcal{B} = \{1, 1 + x, 1 + x + x^2\}$, con el producto interno del ejemplo 7.5

41. **(a)** Calcule los tres primeros polinomios normalizados de Legendre. (Vea el ejemplo 7.8.)

(b) Use el proceso de Gram-Schmidt para encontrar el cuarto polinomio normalizado de Legendre.

42. Si el polinomio de Legendre de grado n se multiplica por un escalar adecuado es posible obtener un polinomio $L_n(x)$ tal que $L_n(1) = 1$.

(a) Encuentre $L_0(x)$, $L_1(x)$, $L_2(x)$ y $L_3(x)$.

(b) Puede demostrarse que $L_n(x)$ satisface la relación de recurrencia

$$L_n(x) = \frac{2n - 1}{n} x L_{n-1}(x) - \frac{n - 1}{n} L_{n-2}(x)$$

para todo $n \geq 2$. Verifique esta recurrencia para $L_2(x)$ y $L_3(x)$. Luego úsela para calcular $L_4(x)$ y $L_5(x)$.

43. Verifique que si W es un subespacio de un espacio V con producto interno y **v** está en V, entonces $\text{perp}_W(\mathbf{v})$ es ortogonal a todo **w** en W.

44. Sean $\mathbf{u}$ y $\mathbf{v}$ vectores en un espacio con producto interno V. Demuestre la Desigualdad de Cauchy-Schwarz para $\mathbf{u} \neq \mathbf{0}$ del modo siguiente:

(a) Sea t un escalar real. Entonces $\langle t\mathbf{u} + \mathbf{v}, t\mathbf{u} + \mathbf{v} \rangle \geq 0$ para todos los valores de t. Expanda esta desigualdad para obtener una desigualdad cuadrática de la forma

$$at^2 + bt + c \geq 0$$

¿Cuáles son a, b y c en términos de $\mathbf{u}$ y $\mathbf{v}$?

(b) Use su conocimiento de las ecuaciones cuadráticas y sus gráficas para obtener una condición sobre a, b y c para la cual sea verdadera la desigualdad en el inciso (a).

(c) Demuestre que, en términos de $\mathbf{u}$ y $\mathbf{v}$, su condición en el inciso (b) es equivalente a la desigualdad de Cauchy-Schwarz.

Exploración

Vectores y matrices con entradas complejas

En este libro se desarrollaron la teoría y las aplicaciones de los espacios vectoriales reales, cuyo ejemplo más básico es $\mathbb{R}^n$. También se exploraron los espacios vectoriales finitos $\mathbb{Z}_p^n$ y sus aplicaciones. El conjunto $\mathbb{C}^n$ de n-tuplos de números complejos también es un espacio vectorial, con los números complejos $\mathbb{C}$ como escalares. Los axiomas de espacio vectorial (sección 6.1) se cumplen todos para $\mathbb{C}^n$, y los conceptos como independencia lineal, base y dimensión se trasladan desde $\mathbb{R}^n$ sin dificultad.

La primera diferencia notable entre $\mathbb{R}^n$ y $\mathbb{C}^n$ está en la definición de producto punto. Si el producto punto se define en $\mathbb{C}^n$ como en $\mathbb{R}^n$, entonces para el vector $\mathbf{v} = \begin{bmatrix} i \\ 1 \end{bmatrix}$ distinto de cero se tiene

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{i^2 + 1^2} = \sqrt{-1 + 1} = \sqrt{0} = 0$$

Ésta claramente es una situación indeseable (un vector distinto de cero cuya longitud es cero) y viola los Teoremas 1.2(d) y 1.3. Ahora el producto punto se generaliza a $\mathbb{C}^n$ en una forma que evita este tipo de dificultad.

Definición Si $\mathbf{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ son vectores en $\mathbb{C}^n$, entonces el ***producto punto complejo*** de $\mathbf{u}$ y $\mathbf{v}$ se define mediante

$$\mathbf{u}\cdot\mathbf{v} = \bar{u}_1 v_1 + \cdots + \bar{u}_n v_n$$

La norma (o longitud) de un vector complejo $\mathbf{v}$ se define como en el caso real: $\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}}$. Del mismo modo, la distancia entre dos vectores complejos $\mathbf{u}$ y $\mathbf{v}$ todavía se define como $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$.

1. Demuestre que, para $\mathbf{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ en $\mathbb{C}^n$, $\|\mathbf{v}\| = \sqrt{|v_1^2| + |v_2^2| + \cdots + |v_n^2|}$.

2. Sean $\mathbf{u} = \begin{bmatrix} i \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 2 - 3i \\ 1 + 5i \end{bmatrix}$. Encuentre

(a) $\mathbf{u} \cdot \mathbf{v}$ (b) $\|\mathbf{u}\|$ (c) $\|\mathbf{v}\|$ (d) $d(\mathbf{u}, \mathbf{v})$ (e) un vector distinto de cero ortogonal a $\mathbf{u}$
(f) un vector distinto de cero ortogonal a $\mathbf{v}$

El producto punto complejo es un ejemplo de la noción más general de un producto interno complejo, que satisface las mismas condiciones que un producto interno real con dos excepciones. El problema 3 ofrece un resumen.

3. Demuestre que el producto punto complejo satisface las siguientes propiedades para todos los vectores $\mathbf{u}$, $\mathbf{v}$ y $\mathbf{w}$ en $\mathbb{C}^n$ y todos los escalares complejos.

(a) $\mathbf{u} \cdot \mathbf{v} = \overline{\mathbf{v} \cdot \mathbf{u}}$
(b) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$
(c) $(c\mathbf{u}) \cdot \mathbf{v} = \bar{c}(\mathbf{u} \cdot \mathbf{v})$ y $\mathbf{u} \cdot (c\mathbf{v}) = c(\mathbf{u} \cdot \mathbf{v})$
(d) $\mathbf{u} \cdot \mathbf{u} \geqslant 0$ y $\mathbf{u} \cdot \mathbf{u} = 0$ si y sólo si $\mathbf{u} = \mathbf{0}$.

Para matrices con entradas complejas, la suma, la multiplicación por escalares complejos, la transpuesta y la multiplicación de matrices se definen todas exactamente como se hizo para las matrices reales en la sección 3.1, y las propiedades algebraicas de dichas operaciones todavía se cumplen. (Vea la sección 3.2.) Del mismo modo, se tiene la noción de inversa y el determinante de una matriz cuadrada compleja tal como en el caso real, y las técnicas y propiedades se trasladan todas al caso complejo. (Vea las secciones 3.3 y 4.2.)

Sin embargo, la noción de transpuesta es menos útil en el caso complejo que en el caso real. La siguiente definición proporciona una alternativa.

Definición Si A es una matriz compleja, entonces la ***transpuesta conjugada*** de A es la matriz A^* definida por

$$A^* = \bar{A}^T$$

En la definición anterior, $\bar{A}$ se refiere a la matriz cuyas entradas son los conjugados complejos de las correspondientes entradas de A; esto es, si $A = [a_{ij}]$, entonces $\bar{A} = [\bar{a}_{ij}]$.

4. Encuentre la transpuesta conjugada A^* de la matriz dada:

(a) $A = \begin{bmatrix} i & 2i \\ -i & 3 \end{bmatrix}$ (b) $A = \begin{bmatrix} 2 & 5 - 2i \\ 5 + 2i & -1 \end{bmatrix}$

(c) $A = \begin{bmatrix} 2 - i & 1 + 3i & -2 \\ 4 & 0 & 3 - 4i \end{bmatrix}$ (d) $A = \begin{bmatrix} 3i & 0 & 1 + i \\ 1 - i & 4 & i \\ 1 + i & 0 & -i \end{bmatrix}$

Las propiedades del conjugado complejo (Apéndice C) se extienden a las matrices, como muestra el siguiente problema.

5. Sean A y B matrices complejas y sea c un escalar complejo. Demuestre las siguientes propiedades:

(a) $\bar{\bar{A}} = A$ (b) $\overline{A + B} = \bar{A} + \bar{B}$
(c) $\overline{cA} = \bar{c}\bar{A}$ (d) $\overline{AB} = \bar{A}\bar{B}$
(e) $(\bar{A})^T = (\overline{A^T})$

Las propiedades del problema 5 pueden usarse para establecer las siguientes propiedades de la transpuesta conjugada, que son análogas a las propiedades de la transpuesta para matrices reales (Teorema 3.4).

6. Sean A y B matrices complejas y sea c un escalar complejo. Demuestre las siguientes propiedades:

(a) $(A^*)^* = A$ (b) $(A + B)^* = A^* + B^*$
(c) $(cA)^* = \bar{c}A^*$ (d) $(AB)^* = B^*A^*$

7. Demuestre que para vectores $\mathbf{u}$ y $\mathbf{v}$ en $\mathbb{C}^n$, el producto punto complejo satisface $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^*\mathbf{v}$. (Este resultado es por lo que el producto punto complejo se definió como se hizo. Proporciona el análogo de la fórmula $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$ para vectores en $\mathbb{R}^n$.)

Para matrices reales, se vio la importancia de las matrices simétricas, en especial en el estudio de la diagonalización. Recuerde que una matriz real A es simétrica si $A^T = A$. Para matrices complejas, la siguiente definición es la generalización correcta.

Las matrices hermitianas reciben su nombre en honor del matemático francés Charles Hermite (1822-1901). Hermite es mejor conocido por su prueba de que el número *e* es trascendental, pero también fue el primero en usar el término *matrices ortogonales* y probó que las matrices simétricas (y hermitianas) tienen eigenvalores reales.

Definición Una matriz compleja cuadrada A se llama ***hermitiana*** si $A^* = A$; es decir, si es igual a su propia transpuesta conjugada.

8. Demuestre que las entradas diagonales de una matriz hermitiana deben ser reales.

9. ¿Cuáles de las siguientes matrices son hermitianas?

(a) $A = \begin{bmatrix} 2 & 1+i \\ 1-i & i \end{bmatrix}$ (b) $A = \begin{bmatrix} -1 & 2-3i \\ 2-3i & 5 \end{bmatrix}$

(c) $A = \begin{bmatrix} -3 & -1+5i \\ 1-5i & 3 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & 1+4i & 3-i \\ 1-4i & 2 & i \\ 3+i & -i & 0 \end{bmatrix}$

(e) $A = \begin{bmatrix} 0 & 3 & 2 \\ -3 & 0 & -1 \\ -2 & 1 & 0 \end{bmatrix}$ (f) $A = \begin{bmatrix} 3 & 0 & -2 \\ 0 & 2 & 1 \\ -2 & 1 & 5 \end{bmatrix}$

10. Demuestre que los eigenvalores de una matriz hermitiana son números reales. [*Sugerencia:* la demostración del Teorema 5.18 puede adaptarse al usar la operación transpuesta conjugada.]

11. Demuestre que si A es una matriz hermitiana, entonces los correspondientes eigenvectores para distintos eigenvalores de A son ortogonales. [*Sugerencia:* adapte la demostración del Teorema 5.19 usando $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^* \mathbf{v}$ en lugar de $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v}$.]

Recuerde que una matriz real cuadrada Q es ortogonal si $Q^{-1} = Q^T$. La siguiente definición brinda el análogo complejo.

Definición Una matriz compleja cuadrada U se llama ***unitaria*** si $U^{-1} = U^*$.

Como en el caso de las matrices ortogonales, en la práctica no es necesario calcular U^{-1} directamente. Sólo es necesario demostrar que $U^*U = I$ para verificar que U es unitaria.

12. ¿Cuáles de las siguientes matrices son unitarias? Para las que lo sean, proporcione sus inversas.

(a) $\begin{bmatrix} i/\sqrt{2} & -i/\sqrt{2} \\ i/\sqrt{2} & i/\sqrt{2} \end{bmatrix}$ (b) $\begin{bmatrix} 1+i & 1+i \\ 1-i & -1+i \end{bmatrix}$

(c) $\begin{bmatrix} 3/5 & -4/5 \\ 4i/5 & 3i/5 \end{bmatrix}$ (d) $\begin{bmatrix} (1+i)/\sqrt{6} & 0 & 2/\sqrt{6} \\ 0 & 1 & 0 \\ (-1-i)/\sqrt{3} & 0 & 1/\sqrt{3} \end{bmatrix}$

Las matrices unitarias se comportan en muchos aspectos como las matrices ortogonales. El siguiente problema ofrece algunas caracterizaciones alternativas de las matrices unitarias.

13. Demuestre que los siguientes enunciados son equivalentes para una matriz compleja cuadrada U:

(a) U es unitaria.
(b) Las columnas de U forman un conjunto ortonormal en $\mathbb{C}^n$ con respecto al producto punto complejo.
(c) Los renglones de U forman un conjunto ortonormal en $\mathbb{C}^n$ con respecto al producto punto complejo.
(d) $\|U\mathbf{x}\| = \|\mathbf{x}\|$ para todo $\mathbf{x}$ en $\mathbb{C}^n$.
(e) $U\mathbf{x}\cdot U\mathbf{y} = \mathbf{x}\cdot\mathbf{y}$ para todo $\mathbf{x}$ y $\mathbf{y}$ en $\mathbb{C}^n$.
[*Sugerencia:* Adapte las demostraciones de los Teoremas 5.4-5.7.]

14. Repita el problema 12, esta vez al aplicar los criterios del inciso (b) o el inciso (c) del problema 13.

La siguiente definición es la generalización natural de la diagonalizabilidad ortogonal de las matrices complejas.

Definición Una matriz compleja cuadrada A se llama ***diagonalizable unitariamente*** si existe una matriz unitaria U y una matriz diagonal D tal que

$$U^*AU = D$$

El proceso de diagonalizar una matriz A de $n \times n$ diagonalizable unitariamente imita el caso real. Las columnas de U deben formar una base ortonormal para $\mathbb{C}^n$ que consiste de los eigenvectores de A. Por tanto, (1) se calculan los eigenvalores de A, (2) se encuentra una base para cada eigenespacio, (3) se garantiza que cada base de eigenespacio consista de vectores ortonormales (con el uso del proceso de Gram-Schmidt y del producto punto complejo, si es necesario), (4) se forma la matriz U cuyas columnas son los eigenvectores ortonormales recién encontrados. Entonces U^*AU será una matriz diagonal D cuyas entradas diagonales son los eigenvalores de A, colocados en el mismo orden que los correspondientes eigenvectores en las columnas de U.

15. En cada uno de los siguientes, encuentre una matriz unitaria U y una matriz diagonal D tal que $U^*AU = D$.

(a) $A = \begin{bmatrix} 2 & i \\ -i & 2 \end{bmatrix}$ (b) $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

(c) $A = \begin{bmatrix} -1 & 1+i \\ 1-i & 0 \end{bmatrix}$ (d) $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 1-i \\ 0 & 1+i & 3 \end{bmatrix}$

Las matrices en (a), (c) y (d) del problema anterior son todas hermitianas. Es evidente que toda matriz hermitiana es diagonalizable unitariamente. (Este es el *teorema espectral complejo*, que puede demostrarse al adaptar la prueba del Teorema 5.20.) En este punto quizá sospeche que lo inverso de este resultado también debe ser verdadero; a saber, que toda matriz diagonalizable unitariamente debe ser hermitiana. Pero, por desgracia, ¡esto es *falso*! (¿Puede ver dónde deja de funcionar el análogo complejo de la prueba del Teorema 5.17?)



Para un contraejemplo específico, tome la matriz del inciso (b) del problema 15. No es hermitiana, pero *sí* es diagonalizable unitariamente.

Es evidente que la caracterización correcta de la diagonalizabilidad unitaria es el siguiente teorema, cuya demostración puede encontrarse en libros de texto más avanzados.

Vea *Linear Algebra with Applications* de S. J. Leon (Upper Saddle River, NJ: Prentice-Hall, 2002).

Una matriz compleja cuadrada A es ***diagonalizable unitariamente*** si y sólo si

$$A^*A = AA^*$$

Una matriz A para la cual $A^*A = AA^*$ se llama ***normal***.

16. Demuestre que toda matriz hermitiana, toda matriz unitaria y toda matriz *antihermitiana* $(A^* = -A)$ es normal. (Note que, en el caso real, este resultado se refiere a matrices simétricas, ortogonales y antisimétricas, respectivamente.)

17. Demuestre que si una matriz cuadrada compleja es diagonalizable unitariamente, entonces debe ser normal.

Desigualdades geométricas y problemas de optimización

Esta exploración introducirá algunas aplicaciones poderosas (y acaso sorprendentes) de varias desigualdades, como la desigualdad de Cauchy-Schwarz. Como verá, ciertos problemas de maximización-minimización (*problemas de optimización*), que usualmente surgen en un curso de cálculo, ¡pueden resolverse sin usar cálculo!

Recuerde que la desigualdad de Cauchy-Schwarz en $\mathbb{R}^n$ afirma que, para todos los vectores $\mathbf{u}$ y $\mathbf{v}$,

$$|\mathbf{u}\cdot\mathbf{v}| \le \|\mathbf{u}\|\|\mathbf{v}\|$$

con igualdad si y sólo si $\mathbf{u}$ y $\mathbf{v}$ son múltiplos escalares uno de otro. Si $\mathbf{u} = [x_1 \ \cdots \ x_n]^T$ y $\mathbf{v} = [y_1 \ \cdots \ y_n]^T$, la desigualdad anterior es equivalente a

$$|x_1y_1 + \cdots + x_ny_n| \le \sqrt{x_1^2 + \cdots + x_n^2}\sqrt{y_1^2 + \cdots + y_n^2}$$

Al elevar al cuadrado ambos lados y usar notación de suma, se tiene

$$\left(\sum_{i=1}^{n} x_iy_i\right)^2 \le \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$$

La igualdad se sostiene si y sólo si existe algún escalar k tal que $y_i = kx_i$ para $i = 1, \ldots, n$.

Comience por usar Cauchy-Schwarz para obtener un caso especial de una de las más útiles de todas las desigualdades.

1. Sean x y y números reales no negativos. Aplique la desigualdad de Cauchy-Schwarz a $\mathbf{u} = \begin{bmatrix} \sqrt{x} \\ \sqrt{y} \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} \sqrt{y} \\ \sqrt{x} \end{bmatrix}$ para demostrar que

$$\sqrt{xy} \leq \frac{x + y}{2} \quad (1)$$

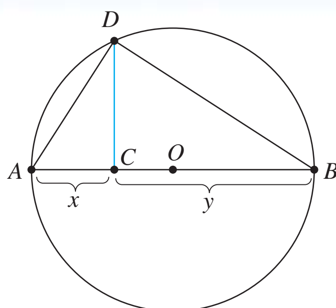


Figura 7.4

con igualdad si y sólo si $x = y$.

2. (a) Demuestre la desigualdad (1) directamente. [*Sugerencia:* eleve al cuadrado ambos lados.] (b) La figura 7.4 muestra un círculo con centro O y diámetro $AB = AC + CB = x + y$. El segmento $\overline{CD}$ es perpendicular a $\overline{AB}$. Demuestre que $CD = \sqrt{xy}$ y use este resultado para deducir la desigualdad (1). [*Sugerencia:* use triángulos semejantes.]

El lado derecho de la desigualdad (1) es la familiar ***media aritmética*** (o *promedio*) de los números x y y. El lado izquierdo muestra la menos familiar ***media geométrica*** de x y y. En concordancia, la desigualdad (1) se conoce como la ***desigualdad media aritmética-media geométrica (MA-MG)***. Se sustenta de manera más general; para n variables no negativas $x_1, \ldots, x_n$, afirma que

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}$$

con igualdad si y sólo si $x_1 = x_2 = \cdots = x_n$.

En palabras: la desigualdad MA-MG dice que la media geométrica de un conjunto de números no negativos siempre es menor o igual a su media aritmética, y las dos son precisamente la misma cuando todos los números son iguales. (Para la demostración general, vea el Apéndice B.)

Ahora se explorará cómo puede aplicarse tal desigualdad a problemas de optimización. He aquí un problema de cálculo típico.

Ejemplo 7.9

Demuestre que, entre todos los rectángulos cuyo perímetro es 100 unidades, el cuadrado tiene el área más grande.

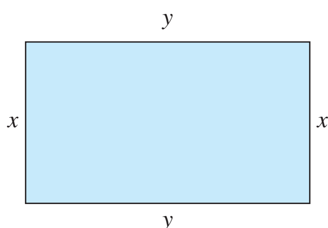


Figura 7.5

Solución Si se hacen x y y las dimensiones del rectángulo (vea la figura 7.5), entonces el área que se quiere maximizar está dada por

$$A = xy$$

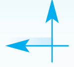
Se indica que el perímetro satisface

$$2x + 2y = 100$$

que es lo mismo que $x + y = 50$. Es posible relacionar xy y $x + y$ usando la desigualdad MA-MG:

$$\sqrt{xy} \leq \frac{x + y}{2} \quad \text{o, de manera equivalente,} \quad xy \leq \tfrac{1}{4}(x + y)^2$$

Dado que $x + y = 50$ es una *constante* (y ésta es la clave), se ve que el valor máximo de $A = xy$ es $50^2/4 = 625$ y ocurre cuando $x = y = 25$.



¡No hay derivada a la vista! ¿No es impresionante? Note que, en este problema de maximización, el paso crucial fue demostrar que el lado derecho de la desigualdad MA-MG era *constante*. En forma similar, es posible aplicar la desigualdad a un problema de *minimización* si puede arreglar que el lado izquierdo sea constante.

Ejemplo 7.10

Demuestre que entre todos los prismas rectangulares con volumen de 8 m^3, el cubo tiene el área superficial mínima.

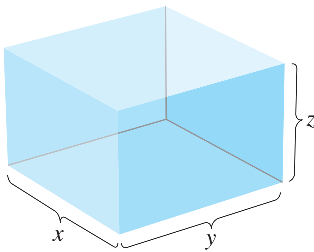


Figura 7.6

Solución Como se muestra en la figura 7.6, si las dimensiones de tal prisma son x, y y z, entonces su volumen está dado por

$$V = xyz$$

Por ende, se tiene $xyz = 8$. El área superficial a minimizar es

$$S = 2xy + 2yz + 2zx$$

Puesto que este es un problema de tres variables, la cosa obvia por intentar es la versión de la desigualdad MA-MG para $n = 3$, a saber

$$\sqrt[3]{xyz} \leq \frac{x + y + z}{3}$$

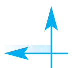
Por desgracia, la expresión para S no aparece aquí. Sin embargo, la desigualdad MA-MG también implica que

$$\begin{aligned} \frac{S}{3} &= \frac{2xy + 2yz + 2zx}{3} \\ &\geq \sqrt[3]{(2xy)(2yz)(2zx)} \\ &= 2\sqrt[3]{(xyz)^2} \\ &= 2\sqrt[3]{64} = 8 \end{aligned}$$

que es equivalente a $S \geq 24$. Por tanto, el valor mínimo de S es 24, y ocurre cuando

$$2xy = 2yz = 2zx$$


(¿Por qué?) Esto implica que $x = y = z = 2$ (es decir, el prisma rectangular es un cubo).



3. Demuestre que entre todos los rectángulos con área de 100 unidades cuadradas, el cuadrado tiene el perímetro más pequeño.

4. ¿Cuál es el valor mínimo de $f(x) = x + \dfrac{1}{x}$ para $x > 0$?

5. Una caja de cartón, con una base cuadrada y parte superior abierta, se construirá a partir de un cuadrado de cartón de 10 cm de lado al cortar cuatro cuadrados en las esquinas y doblar los lados. ¿Cuáles deben ser las dimensiones de la caja con la finalidad de hacer el volumen encerrado tan grande como sea posible?

6. Encuentre el valor mínimo de $f(x, y, z) = (x + y)(y + z)(z + x)$ si x, y y z son números reales positivos tales que $xyz = 1$.

7. Para $x > y > 0$, encuentre el valor mínimo de $x + \dfrac{8}{y(x - y)}$. [*Sugerencia:* puede ser útil una sustitución.]

La desigualdad de Cauchy-Schwarz puede aplicarse a problemas similares, como ilustra el siguiente ejemplo.

Ejemplo 7.11

Encuentre el valor máximo de la función $f(x, y, z) = 3x + y + 2z$ sujeta a la restricción $x^2 + y^2 + z^2 = 1$. ¿Dónde ocurre el valor máximo?

Solución Este tipo de problema por lo general se maneja con técnicas que se estudian en un curso de cálculo de varias variables. Aquí se muestra cómo usar la desigualdad de Cauchy-Schwarz. La función $3x + y + 2z$ tiene la forma de un producto punto, así que se hace

$$\mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix} \quad \text{y} \quad \mathbf{v} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Entonces la forma en componentes de la Desigualdad de Cauchy-Schwarz produce

$$(3x + y + 2z)^2 \leq (3^2 + 1^2 + 2^2)(x^2 + y^2 + z^2) = 14$$

Por tanto, el valor máximo de la función es $\sqrt{14}$, y ocurre cuando

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = k \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}$$

Por tanto, $x = 3k$, $y = k$ y $z = 2k$, de modo que $3(3k) + k + 2(2k) = \sqrt{14}$. Se tiene que $k = 1/\sqrt{14}$, y en consecuencia

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3/\sqrt{14} \\ 1/\sqrt{14} \\ 2/\sqrt{14} \end{bmatrix}$$

8. Encuentre el valor máximo de $f(x,y,z) = x + 2y + 4z$ sujeta a $x^2 + 2y^2 + z^2 = 1$.

9. Encuentre el valor mínimo de $f(x, y, z) = x^2 + y^2 + \dfrac{z^2}{2}$ sujeta a $x + y + z = 10$.

10. Encuentre el valor máximo de sen θ + cos θ.

11. Encuentre el punto sobre la recta $x + 2y = 5$ que esté más cerca del origen.

Existen muchas otras desigualdades que pueden usarse para resolver problemas de optimización. La ***media cuadrática*** de los números $x_1, \ldots, x_n$ se define como

$$\sqrt{\frac{x_1^2 + \cdots + x_n^2}{n}}$$

Si $x_1, \ldots, x_n$ son distintos de cero, su ***media armónica*** está dada por

$$\frac{n}{1/x_1 + 1/x_2 + \cdots + 1/x_n}$$

Es evidente que las medias cuadrática, aritmética, geométrica y armónica están todas relacionadas.

12. Sean x y y números reales positivos. Demuestre que

$$\sqrt{\frac{x^2 + y^2}{2}} \geq \frac{x + y}{2} \geq \sqrt{xy} \geq \frac{2}{1/x + 1/y}$$

con igualdad si y sólo si $x = y$. (La desigualdad en el medio es justo MA-MG, así que sólo es necesario establecer la primera y tercera desigualdades.)

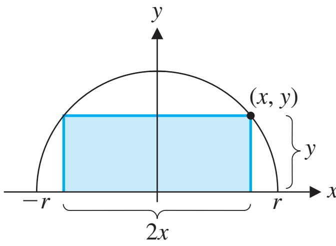


Figura 7.7

13. Encuentre el área del rectángulo más grande que puede inscribirse en un semicírculo de radio r (figura 7.7).

14. Encuentre el valor mínimo de la función

$$f(x, y) = \frac{(x + y)^2}{xy}$$

para $x, y > 0$. [*Sugerencia:* $(x + y)^2/xy = (x + y)(1/x + 1/y)$.]

15. Sean x y y números reales positivos con $x + y = 1$. Demuestre que el valor mínimo de

$$f(x, y) = \left(x + \frac{1}{x}\right)^2 + \left(y + \frac{1}{y}\right)^2$$

es $\frac{25}{2}$, y determine los valores de x y y para los cuales ocurre.

7.2 Normas y funciones de distancia

En la última sección vio que es posible definir la longitud y la distancia en un espacio con producto interno. Como verá dentro de poco, también existen algunas versiones de estos dos conceptos que no están definidos en términos de un producto interno.

Para comenzar, es necesario especificar las propiedades que es deseable tenga una "función de longitud". La siguiente definición hace esto, y se usa como su base el Teorema 1.3 y la desigualdad del triángulo.

Definición Una ***norma*** en un espacio vectorial V es un mapeo que asocia con cada vector $\mathbf{v}$ un número real $\|\mathbf{v}\|$, llamado la ***norma*** de $\mathbf{v}$, tal que se satisfacen las siguientes propiedades para todos los vectores $\mathbf{u}$ y $\mathbf{v}$ y todos los escalares c:

1. $\|\mathbf{v}\| \geq 0$ y $\|\mathbf{v}\| = 0$ si y sólo si $\mathbf{v} = \mathbf{0}$.
2. $\|c\mathbf{v}\| = |c|\|\mathbf{v}\|$
3. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

Un espacio vectorial con una norma se llama ***espacio lineal normado***.

Ejemplo 7.12 Demuestre que, en un espacio con producto interno, $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$ se define una norma.

Solución Claramente, $\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} \geq 0$. Más aún,

$$\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = 0 \Leftrightarrow \langle \mathbf{v}, \mathbf{v} \rangle = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$$

por la definición de producto interno. Esto demuestra la propiedad (1).

Para la propiedad (2), sólo es necesario notar que

$$\|c\mathbf{v}\| = \sqrt{\langle c\mathbf{v}, c\mathbf{v} \rangle} = \sqrt{c^2\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{c^2}\sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = |c|\|\mathbf{v}\|$$

La propiedad (3) es justo la desigualdad del triángulo, lo que se verificó en el Teorema 7.4.

Ahora se observarán algunos ejemplos de normas que no están definidas en términos de un producto interno. El ejemplo 7.13 es la generalización matemática a $\mathbb{R}^n$ de la norma de taxi que se exploró en la introducción de este capítulo.

Ejemplo 7.13 La ***norma suma*** $\|\mathbf{v}\|_s$ de un vector $\mathbf{v}$ en $\mathbb{R}^n$ es la suma de los valores absolutos de sus componentes. Esto es, si $\mathbf{v} = [v_1 \; \cdots \; v_n]^T$, entonces

$$\|\mathbf{v}\|_s = |v_1| + \cdots + |v_n|$$

Demuestre que la norma suma es una norma.

Solución Claramente, $\|\mathbf{v}\|_s = |v_1| + \cdots + |v_n| \geq 0$, y la única forma de lograr la igualdad es si $|v_1| = \cdots = |v_n| = 0$. Pero esto es así si y sólo si $v_1 = \cdots = v_n = 0$ o, de manera equivalente, $\mathbf{v} = \mathbf{0}$, lo que demuestra la propiedad (1). Para la propiedad (2), se ve que $c\mathbf{v} = [cv_1 \;\; \cdots \;\; cv_n]^T$, de modo que

$$\|c\mathbf{v}\|_s = |cv_1| + \cdots + |cv_n| = |c|(|v_1| + \cdots + |v_n|) = |c|\|\mathbf{v}\|_s$$

Finalmente, se sustenta la desigualdad del triángulo, porque si $\mathbf{u} = [u_1 \ \cdots \ u_n]^T$, entonces

$$\begin{aligned}\|\mathbf{u} + \mathbf{v}\|_s &= |u_1 + v_1| + \cdots + |u_n + v_n| \\ &\leq (|u_1| + |v_1|) + \cdots + (|u_n| + |v_n|) \\ &= (|u_1| + \cdots + |u_n|) + (|v_1| + \cdots + |v_n|) = \|\mathbf{u}\|_s + \|\mathbf{v}\|_s\end{aligned}$$

La norma suma también se conoce como la **1-*norma*** y con frecuencia se denota mediante $\|\mathbf{v}\|_1$. Sobre $\mathbb{R}^2$ es lo mismo que la norma de taxi. Como demuestra el ejemplo 7.13, es posible tener varias normas en el mismo espacio vectorial. El ejemplo 7.14 ilustra otra norma sobre $\mathbb{R}^n$.

Ejemplo 7.14

La ***norma máx*** $\|\mathbf{v}\|_m$ de un vector $\mathbf{v}$ en $\mathbb{R}^n$ es el número más grande entre los valores absolutos de sus componentes. Esto es, si $\mathbf{v} = [v_1 \ \cdots \ v_n]^T$, entonces

$$\|\mathbf{v}\|_m = \text{máx}\{|v_1|, \ldots, |v_n|\}$$

Demuestre que la norma máx es una norma.

Solución Nuevamente, es claro que $\|\mathbf{v}\|_m \geq 0$. Si $\|\mathbf{v}\|_m = 0$, entonces el mayor de $|v_1|, \ldots, |v_n|$ es cero, y por tanto todos lo son. En consecuencia, $v_1 = \cdots = v_n = 0$, de modo que $\mathbf{v} = \mathbf{0}$. Esto verifica la propiedad (1). A continuación, observe que para cualquier escalar c,

$$\|c\mathbf{v}\|_m = \text{máx}\{|cv_1|, \ldots, |cv_n|\} = |c|\text{máx}\{|v_1|, \ldots, |v_n|\} = |c|\|\mathbf{v}\|_m$$

Finalmente, para $\mathbf{u} = [u_1 \ \cdots \ u_n]^T$, se tiene

$$\begin{aligned}\|\mathbf{u} + \mathbf{v}\|_m &= \text{máx}\{|u_1 + v_1|, \ldots, |u_n + v_n|\} \\ &\leq \text{máx}\{|u_1| + |v_1|, \ldots, |u_n| + |v_n|\} \\ &\leq \text{máx}\{|u_1|, \ldots, |u_n|\} + \text{máx}\{|v_1|, \ldots, |v_n|\} = \|\mathbf{u}\|_m + \|\mathbf{v}\|_m\end{aligned}$$

(¿Por qué la segunda desigualdad es verdadera?) Esto verifica la desigualdad del triángulo.

La norma máx también se conoce como la **∞-*norma*** o ***norma uniforme*** y con frecuencia se denota mediante $\|\mathbf{v}\|_\infty$. En general, es posible definir una norma $\|\mathbf{v}\|_p$ sobre $\mathbb{R}^n$ mediante

$$\|\mathbf{v}\|_p = (|v_1|^p + \cdots + |v_n|^p)^{1/p}$$

para cualquier número real $p \geq 1$. Para $p = 1$, $\|\mathbf{v}\|_1 = \|\mathbf{v}\|_s$, lo que justifica el término 1-norma. Para $p = 2$,

$$\|\mathbf{v}\|_2 = (|v_1|^2 + \cdots + |v_n|^2)^{1/2} = \sqrt{v_1^2 + \cdots + v_n^2}$$

que es justo la norma familiar sobre $\mathbb{R}^n$ que se obtiene a partir del producto punto. Llamada **2-*norma*** o ***norma euclidiana***, con frecuencia se denota $\|\mathbf{v}\|_E$. Conforme p se hace grande, puede demostrarse usando cálculo que $\|\mathbf{v}\|_p$ tiende a la norma máx $\|\mathbf{v}\|_m$. Esto justifica el uso de la notación alternativa $\|\mathbf{v}\|_\infty$ para esta norma.

Ejemplo 7.15

Para un vector $\mathbf{v}$ en $\mathbb{Z}_2^n$, defina $\|\mathbf{v}\|_H$ como $w(\mathbf{v})$, el peso de $\mathbf{v}$. Demuestre que esto define una norma.

Solución Ciertamente, $\|\mathbf{v}\|_H = w(\mathbf{v}) \geq 0$ y el único vector cuyo peso es cero es el vector cero. Por tanto, la propiedad (1) es verdadera. Dado que los únicos candidatos para el escalar c son 0 y 1, la propiedad (2) es inmediata.

Para verificar la desigualdad del triángulo, observe primero que si $\mathbf{u}$ y $\mathbf{v}$ son vectores en $\mathbb{Z}_2^n$, entonces $w(\mathbf{u} + \mathbf{v})$ cuenta el número de lugares en los que difieren $\mathbf{u}$ y $\mathbf{v}$. [Por ejemplo, si

$$\mathbf{u} = [1 \quad 1 \quad 0 \quad 1 \quad 0]^T \quad \text{y} \quad \mathbf{v} = [0 \quad 1 \quad 1 \quad 1 \quad 1]^T$$

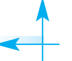
entonces $\mathbf{u} + \mathbf{v} = [1 \quad 0 \quad 1 \quad 0 \quad 1]^T$, de modo que $w(\mathbf{u} + \mathbf{v}) = 3$, en concordancia con el hecho de que $\mathbf{u}$ y $\mathbf{v}$ difieren exactamente en tres posiciones.] Suponga que tanto $\mathbf{u}$ como $\mathbf{v}$ tienen ceros en las posiciones n_0 y números 1 en las posiciones n_1, $\mathbf{u}$ tiene 0 y $\mathbf{v}$ tiene un 1 en las pocisiones n_{01}, y $\mathbf{u}$ tiene un 1 y $\mathbf{v}$ tiene un 0 en las posiciones n_{10}. (En el ejemplo anterior, $n_0 = 0$, $n_1 = 2$, $n_{01} = 2$, y $n_{10} = 1$.) Ahora

$$w(\mathbf{u}) = n_1 + n_{10}, \quad w(\mathbf{v}) = n_1 + n_{01}, \quad \text{y} \quad w(\mathbf{u} + \mathbf{v}) = n_{10} + n_{01}$$

Por tanto,

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|_H = w(\mathbf{u} + \mathbf{v}) &= n_{10} + n_{01} \\ &= (n_1 + n_{10}) + (n_1 + n_{01}) - 2n_1 \\ &\leq (n_1 + n_{10}) + (n_1 + n_{01}) \\ &= w(\mathbf{u}) + w(\mathbf{v}) = \|\mathbf{u}\|_H + \|\mathbf{v}\|_H \end{aligned}$$

La norma $\|\mathbf{v}\|_H$ se llama ***norma de Hamming***.



Funciones de distancia

Para cualquier norma, se puede definir una función de distancia tal como se hizo en la última sección, a saber,

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Ejemplo 7.16

Sean $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Calcule $d(\mathbf{u}, \mathbf{v})$ en relación con (a) la norma euclidiana, (b) la norma suma y (c) la norma máx.

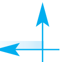
Solución Cada cálculo requiere saber que $\mathbf{u} - \mathbf{v} = \begin{bmatrix} 4 \\ -3 \end{bmatrix}$.

(a) Como ahora es muy familiar,

$$d_E(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|_E = \sqrt{4^2 + (-3)^2} = \sqrt{25} = 5$$

(b) $d_s(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|_s = |4| + |-3| = 7$

(c) $d_m(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|_m = \text{máx}\{|4|, |-3|\} = 4$



La función de distancia sobre $\mathbb{Z}_2^n$ determinada por la norma de Hamming se conoce como ***distancia de Hamming***. En la sección 7.5 se explorará su uso en los códigos de corrección de errores. El ejemplo 7.17 ofrece una ilustración de la distancia de Hamming.

Ejemplo 7.17 Encuentre la distancia de Hamming entre

$$\mathbf{u} = [1 \quad 1 \quad 0 \quad 1 \quad 0]^T \quad \text{y} \quad \mathbf{v} = [0 \quad 1 \quad 1 \quad 1 \quad 1]^T$$

Solución Dado que se trabaja sobre $\mathbb{Z}_2$, $\mathbf{u} - \mathbf{v} = \mathbf{u} + \mathbf{v}$. Pero

$$d_H(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} + \mathbf{v}\|_H = w(\mathbf{u} + \mathbf{v})$$

Como se hizo notar en el ejemplo 7.15, este es justo el número de posiciones en las que difieren **u** y **v**. Los vectores dados son los mismos que se usaron en dicho ejemplo; en consecuencia, el cálculo es exactamente el mismo. Por tanto, $d_H(\mathbf{u}, \mathbf{v}) = 3$.

El Teorema 7.5 resume las propiedades más importantes de una función de distancia.

Teorema 7.5 Sea d una función de distancia definida sobre un espacio lineal normado V. Las siguientes propiedades se cumplen para todos los vectores **u**, **v** y **w** en V:

a. $d(\mathbf{u}, \mathbf{v}) \geq 0$ y $d(\mathbf{u}, \mathbf{v}) = 0$ si y sólo si $\mathbf{u} = \mathbf{v}$.
b. $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$
c. $d(\mathbf{u}, \mathbf{w}) \leq d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w})$

Demostración (a) Al usar la propiedad (1) de la definición de una norma, es fácil comprobar que $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| \geq 0$, donde la igualdad se sostiene si y sólo si $\mathbf{u} - \mathbf{v} = \mathbf{0}$ o, de manera equivalente, $\mathbf{u} = \mathbf{v}$.

(b) En el ejercicio 19 se le pide demostrar la propiedad (b).

(c) Aplique la desigualdad del triángulo para obtener

$$\begin{aligned} d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v}, \mathbf{w}) &= \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{v} - \mathbf{w}\| \\ &\geq \|(\mathbf{u} - \mathbf{v}) + (\mathbf{v} - \mathbf{w})\| \\ &= \|\mathbf{u} - \mathbf{w}\| = d(\mathbf{u}, \mathbf{w}) \end{aligned}$$

A una función d que satisfaga las tres propiedades del Teorema 7.5 también se le llama ***métrica*** y a un espacio vectorial que posea tal función se le llama ***espacio métrico***. Ambos son muy importantes en muchas ramas de las matemáticas y se estudian con detalles en cursos más avanzados.

Normas matriciales

Es posible definir normas para matrices exactamente como se definieron las normas para los vectores en $\mathbb{R}^n$. Después de todo, el espacio vectorial M_{mn} de todas las matrices de $m \times n$ es isomórfico a $\mathbb{R}^{mn}$, de modo que no es difícil hacerlo. Desde luego, las propiedades (1), (2) y (3) de una norma también se cumplen en el escenario de las matrices. Es evidente que, para matrices, las normas que son más útiles satisfacen una propiedad adicional. (La atención se centrará en matrices cuadradas, pero es posible generalizar todo a matrices arbitrarias).

Definición Una ***norma matricial*** sobre M_{nn} es un mapeo que asocia con cada matriz A de $n \times n$ un número real $\|A\|$, llamado ***norma*** de A, tal que todas las matrices A y B de $n \times n$ y todos los escalares c satisfacen las siguientes propiedades.

1. $\|A\| \geq 0$ y $\|A\| = 0$ si y sólo si $A = O$.
2. $\|cA\| = \|A\|$
3. $\|A + B\| \leq \|A\| + \|B\|$
4. $\|AB\| \leq \|A\|\|B\|$

Se dice que una norma matricial sobre M_{nn} es ***compatible*** con una norma vectorial $\|\mathbf{x}\|$ sobre $\mathbb{R}^n$ si, para toda matriz A de $n \times n$ y todos los vectores $\mathbf{x}$ en $\mathbb{R}^n$, se tiene

$$\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|$$

Ejemplo 7.18 La ***norma de Frobenius*** $\|A\|_F$ de una matriz A se obtiene al extender las entradas de la matriz en un vector y luego tomar la norma euclidiana. En otras palabras: $\|A\|_F$ es justo la raíz cuadrada de la suma de los cuadrados de las entradas de A. De este modo, si $A=[a_{ij}]$, entonces

$$\|A\|_F = \sqrt{\sum_{i,j=1}^{n} a_{ij}^2}$$

(a) Encuentre la norma de Frobenius de

$$A = \begin{bmatrix} 3 & -1 \\ 2 & 4 \end{bmatrix}$$

(b) Demuestre que la norma de Frobenius es compatible con la norma euclidiana.

(c) Demuestre que la norma de Frobenius es una norma matricial.

Solución (a) $\|A\|_F = \sqrt{3^2 + (-1)^2 + 2^2 + 4^2} = \sqrt{30}$

Antes de continuar, observe que si $\mathbf{A}_1 = [3 \quad -1]$ y $\mathbf{A}_2 = [2 \quad 4]$ son los vectores renglón de A, entonces $\|\mathbf{A}_1\|_E = \sqrt{3^2 + (-1)^2}$ y $\|\mathbf{A}_2\|_E = \sqrt{2^2 + 4^2}$. Por tanto,

$$\|A\|_F = \sqrt{\|\mathbf{A}_1\|_E^2 + \|\mathbf{A}_2\|_E^2}$$

De igual modo, si $\mathbf{a}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ y $\mathbf{a}_2 = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$ son los vectores columna de A, entonces

$$\|A\|_F = \sqrt{\|\mathbf{a}_1\|_E^2 + \|\mathbf{a}_2\|_E^2}$$

Es fácil ver que estos hechos se extienden a matrices de $n \times n$ en general. Estas observaciones se usarán para resolver los incisos (b) y (c).

(b) Escriba

$$A = \begin{bmatrix} \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_n \end{bmatrix}$$

Entonces

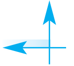
$$\begin{aligned}
\|A\mathbf{x}\|_E &= \left\| \begin{bmatrix} A_1\mathbf{x} \\ \vdots \\ A_n\mathbf{x} \end{bmatrix} \right\|_E \\
&= \sqrt{\|\mathbf{A}_1\mathbf{x}\|_E^2 + \cdots + \|\mathbf{A}_n\mathbf{x}\|_E^2} \\
&\le \sqrt{\|\mathbf{A}_1\|_E^2\|\mathbf{x}\|_E^2 + \cdots + \mathbf{A}_n\|_E^2\|\mathbf{x}\|_E^2} \\
&= \left(\sqrt{\|\mathbf{A}_1\|_{\mathbf{E}}^2 + \cdots + \|\mathbf{A}_n\|_{\mathbf{E}}^2}\right)\|\mathbf{x}\|_{\mathbf{E}} \\
&= \|\mathbf{A}\|_F\|\mathbf{x}\|_E
\end{aligned}$$

donde la desigualdad surge de la desigualdad de Cauchy-Schwarz aplicada a los productos punto de los vectores renglón $\mathbf{A}_i$ con el vector columna $\mathbf{x}$.  (¿Ve cómo se aplicó Cauchy-Schwarz?) Por tanto, la norma de Frobenius es compatible con la norma euclidiana.

(c) Sea $\mathbf{b}_i$ la i-ésima columna de B. Al usar la representación matriz-columna del producto AB, se tiene

$$\begin{aligned}
\|AB\|_F &= \|[A\mathbf{b}_1 \cdots A\mathbf{b}_n]\|_F \\
&= \sqrt{\|A\mathbf{b}_1\|_E^2 + \cdots + \|A\mathbf{b}_n\|_E^2} \\
&\le \sqrt{\|A\|_F^2\|\mathbf{b}_1\|_E^2 + \cdots + \|A\|_F^2\|\mathbf{b}_n\|_E^2} \qquad \text{por el inciso (b)} \\
&= \|A\|_F\sqrt{\|\mathbf{b}_1\|_E^2 + \cdots + \|\mathbf{b}_n\|_E^2} \\
&= \|A\|_F\|B\|_F
\end{aligned}$$

lo que demuestra la propiedad (4) de la definición de una norma matricial. Las propiedades (1) a (3) son verdaderas, pues la norma de Frobenius se obtiene de la norma euclidiana, que satisface dichas propiedades. Por tanto, la norma de Frobenius es una norma matricial.



Para muchas aplicaciones, la norma matricial de Frobenius no es la mejor (o la más sencilla) de usar. Los tipos más útiles de normas matriciales surgen de considerar el efecto de la transformación matricial correspondiente en la matriz cuadrada A. Esta transformación mapea un vector $\mathbf{x}$ en $A\mathbf{x}$. Una forma de medir el "tamaño" de A es comparar $\|\mathbf{x}\|$ y $\|A\mathbf{x}\|$ usando cualquier norma (vectorial) conveniente. Piense por adelantado. Cualquiera que sea la definición de $\|A\|$ a la que se llegue, se sabe que uno quiere que sea compatible con la norma vectorial que se utiliza; esto es, se necesitará

$$\|A\mathbf{x}\| \le \|A\|\|\mathbf{x}\| \quad \text{o} \quad \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \le \|A\| \quad \text{para} \quad \mathbf{x} \ne \mathbf{0}$$

La expresión $\dfrac{\|A\mathbf{x}\|}{\|\mathbf{x}\|}$ mide la "capacidad de estiramiento" de A. Si se normaliza cada vector $\mathbf{x}$ distinto de cero al dividirlo por su norma, se obtienen vectores unitarios $\hat{\mathbf{x}} = \dfrac{1}{\|\mathbf{x}\|}\mathbf{x}$

y por tanto

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{1}{\|\mathbf{x}\|}\|A\mathbf{x}\| = \left\|\frac{1}{\|\mathbf{x}\|}(A\mathbf{x})\right\| = \left\|A\left(\frac{1}{\|\mathbf{x}\|}\mathbf{x}\right)\right\| = \|A\hat{\mathbf{x}}\|$$

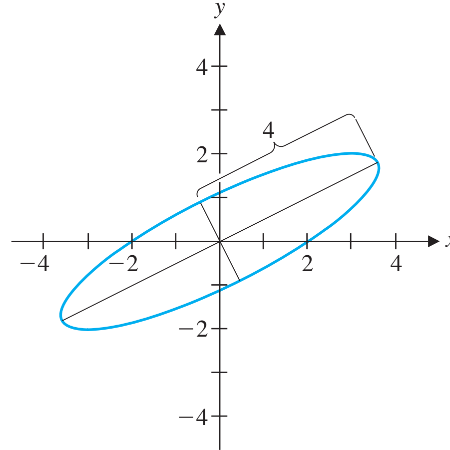


Figura 7.8

Si $\mathbf{x}$ varía sobre todos los vectores distintos de cero en $\mathbb{R}^n$, entonces $\hat{\mathbf{x}}$ varía sobre todos los vectores *unitarios* (esto es, la esfera unitaria) y el conjunto de todos los vectores $A\hat{\mathbf{x}}$ determina alguna curva en $\mathbb{R}^n$. Por ejemplo, la figura 7.8 muestra cómo la matriz $A = \begin{bmatrix} 3 & 2 \\ 2 & 0 \end{bmatrix}$ afecta a la circunferencia unitaria en $\mathbb{R}^2$, mapeándola en una elipse. Con la norma euclidiana, el valor máximo de $\|A\hat{\mathbf{x}}\|$ es claramente justo la mitad de la longitud del eje principal; en este caso, 4 unidades. Esto se expresa al escribir $\max_{\|\hat{\mathbf{x}}\|=1} \|A\hat{\mathbf{x}}\| = 4$.

En la sección 7.4 se verá que este no es un fenómeno aislado. Esto es:

$$\max_{\mathbf{x}\neq\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \max_{\|\hat{\mathbf{x}}\|=1} \|A\hat{\mathbf{x}}\|$$

siempre existe, y hay un vector unitario particular $\mathbf{y}$ para el cual $\|A\mathbf{y}\|$ *es* máximo. Ahora se demostrará que $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$ define una norma matricial.

Teorema 7.6 Si $\|\mathbf{x}\|$ es una norma vectorial sobre $\mathbb{R}^n$, entonces $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$ define una norma matricial sobre M_{nn} que es compatible con la norma vectorial que lo induce.

Demostración (1) Ciertamente, $\|A\mathbf{x}\| \geq 0$ para todos los vectores $\mathbf{x}$, de modo que, en particular, esta desigualdad es verdadera si $\|\mathbf{x}\| = 1$. Por tanto, también $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| \geq 0$. Si $\|A\| = 0$, entonces se debe tener $\|A\mathbf{x}\| = 0$ (y, en consecuencia, $A\mathbf{x} = \mathbf{0}$) para todo $\mathbf{x}$ con $\|\mathbf{x}\| = 1$. En particular, $A\mathbf{e}_i = \mathbf{0}$ para cada uno de los vectores base estándar $\mathbf{e}_i$ en $\mathbb{R}^n$. Pero $A\mathbf{e}_i$ es justo la i-ésima columna de A, de modo que se debe tener $A = O$. Por el contrario, si $A = O$, es claro que $\|A\| = 0$. (¿Por qué?)

(2) Sea c un escalar. Entonces

$$\|cA\| = \max_{\|\mathbf{x}\|=1} \|cA\mathbf{x}\| = \max_{\|\mathbf{x}\|=1} |c|\,\|A\mathbf{x}\| = |c| \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\| = |c|\,\|A\|$$

(3) Sea B una matriz de $n \times n$ y sea $\mathbf{y}$ un vector unitario para el cual

$$\|A + B\| = \max_{\|\mathbf{x}\|=1} \|(A + B)\mathbf{x}\| = \|(A + B)\mathbf{y}\|$$

Entonces

$$\begin{aligned}\|A + B\| &= \|(A + B)\mathbf{y}\| \\ &= \|A\mathbf{y} + B\mathbf{y}\| \\ &\le \|A\mathbf{y}\| + \|B\mathbf{y}\| \\ &\le \|A\| + \|B\|\end{aligned}$$



(¿De dónde proviene la segunda desigualdad?) A continuación se demuestra que la definición es compatible con la norma vectorial [propiedad (5)] y luego se usa este hecho para completar la demostración de que se tiene una norma matricial.

(5) Si $\mathbf{x} = \mathbf{0}$, entonces la desigualdad $\|A\mathbf{x}\| \le \|A\|\|\mathbf{x}\|$ es verdadera, pues ambos lados son cero. Si $\mathbf{x} \ne \mathbf{0}$, entonces, a partir de los comentarios que anteceden a este teorema,

$$\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \le \max_{\mathbf{x}\ne\mathbf{0}} \frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} = \|A\|$$

Por tanto, $\|A\mathbf{x}\| \le \|A\|\|\mathbf{x}\|$.

(4) Sea $\mathbf{z}$ un vector unitario tal que $\|AB\| = \max_{\|\mathbf{x}\|=1} \|(AB)\mathbf{x}\| = \|AB\mathbf{z}\|$. Entonces

$$\begin{aligned}\|AB\| &= \|AB\mathbf{z}\| \\ &= \|A(B\mathbf{z})\| \\ &\le \|A\|\|B\mathbf{z}\| && \text{por la propiedad (5)} \\ &\le \|A\|\|B\|\|\mathbf{z}\| && \text{por la propiedad (5)} \\ &= \|A\|\|B\|\end{aligned}$$

Esto completa la demostración de que $\|A\| = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$ define una norma matricial sobre M_{nn} que es compatible con la norma vectorial que lo induce.

Definición La norma matricial $\|A\|$ en el Teorema 7.6 se llama ***operador norma*** inducida por la norma vectorial $\|\mathbf{x}\|$.

El término *operador norma* refleja el hecho de que una transformación matricial que surge de una matriz cuadrada también se llama *operador lineal*. Por tanto, esta norma es una medida de la capacidad de estiramiento de un operador lineal.

Los tres operadores para normas de uso más común son los inducidos por la norma suma, la norma euclidiana y la norma máx, a saber,

$$\|A\|_1 = \max_{\|\mathbf{x}\|_s=1} \|A\mathbf{x}\|_s, \quad \|A\|_2 = \max_{\|\mathbf{x}\|_E=1} \|A\mathbf{x}\|_E, \quad \|A\|_\infty = \max_{\|\mathbf{x}\|_m=1} \|A\mathbf{x}\|_m$$

respectivamente. La primera y la última evidencian tener fórmulas especialmente agradables que las hacen muy fáciles de calcular.

Teorema 7.7

Sea A una matriz de $n \times n$ con vectores columna $\mathbf{a}_i$ y vectores renglón $\mathbf{A}_i$ para $i = 1, \ldots, n$.

a. $\|A\|_1 = \max_{j=1,\ldots,n} \{\|\mathbf{a}_j\|_s\} = \max_{j=1,\ldots,n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\}$

b. $\|A\|_\infty = \max_{i=1,\ldots,n} \{\|\mathbf{A}_i\|_s\} = \max_{i=1,\ldots,n} \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}$

En otras palabras, $\|A\|_1$ es la suma columna absoluta más grande, y $\|A\|_\infty$ es la suma renglón absoluta más grande. Antes de probar el teorema, observe un ejemplo para ver cuán fácil es usarlo.

Ejemplo 7.19

Sea

$$A = \begin{bmatrix} 1 & -3 & 2 \\ 4 & -1 & -2 \\ -5 & 1 & 3 \end{bmatrix}$$

Encuentre $\|A\|_1$ y $\|A\|_\infty$.

Solución Claramente, la suma absoluta de columna más grande está en la primera columna, de modo que

$$\|A\|_1 = \|\mathbf{a}_1\|_s = |1| + |4| + |-5| = 10$$

El tercer renglón tiene la suma absoluta de renglón más grande, de modo que

$$\|A\|_\infty = \|\mathbf{A}_3\|_s = |-5| + |1| + |3| = 9$$

Con referencia a la definición $\|A\|_1 = \max_{\|\mathbf{x}\|_s=1} \|A\mathbf{x}\|_s$, se ve que el valor máximo de 10 en realidad se logra cuando se toma $\mathbf{x} = \mathbf{e}_1$, pues entonces

$$\|\mathbf{A}\mathbf{e}_1\|_s = \|\mathbf{a}_1\|_s = 10 = \|A\|_1$$

Para $\|A\|_\infty = \max_{\|\mathbf{x}\|_m=1} \|A\mathbf{x}\|_m$, si se toma

$$\mathbf{x} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$$

se obtiene

$$\|A\mathbf{x}\|_m = \left\| \begin{bmatrix} 1 & -3 & 2 \\ 4 & -1 & -2 \\ -5 & 1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\|_m = \left\| \begin{bmatrix} -2 \\ -7 \\ 9 \end{bmatrix} \right\|_m$$

$$= \max\{|-2|, |-7|, |9|\} = 9 = \|A\|_\infty$$

Estas observaciones se usarán para demostrar el Teorema 7.7.

Demostración del Teorema 7.7 La estrategia es la misma que en el caso de la suma de columna y de la suma de renglón. Si M representa el valor máximo, se demuestra que $\|\mathbf{Ax}\| \leq M$ para todos los vectores unitarios $\mathbf{x}$. Entonces se encuentra un vector unitario específico $\mathbf{x}$ para el cual ocurre la igualdad. Es importante recordar que, para la propiedad (a), la norma vectorial es la norma suma, mientras que para la propiedad (b) es la norma máx.

(a) Para demostar (a), sea $M = \max_{j=1,\ldots,n} \{\|\mathbf{a}_j\|_s\}$, la suma columna absoluta máxima y sea $\|\mathbf{x}\|_s = 1$. Entonces $|x_1| + \cdots + |x_n| = 1$, de modo que

$$\begin{aligned}\|\mathbf{A}\mathbf{x}\|_s &= \|x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n\|_s \\ &\leq |x_1|\|\mathbf{a}_1\|_s + \cdots + |x_n|\|\mathbf{a}_n\|_s \\ &\leq |x_1|M + \cdots + |x_n|M \\ &= (|x_1| + \cdots + |x_n|)M = 1 \cdot M = M\end{aligned}$$

Si la suma columna absoluta máxima ocurre en la columna k, entonces con $\mathbf{x} = \mathbf{e}_k$ se obtiene

$$\|\mathbf{Ae}_k\|_s = \|\mathbf{a}_k\|_s = M$$

Por tanto, $\|A\|_1 = \max_{\|\mathbf{x}\|_s=1} \|A\mathbf{x}\|_s = M = \max_{j=1,\ldots,n} \{\|\mathbf{a}_j\|_s\}$, como se requiere.

(b) La demostración de la propiedad (b) se deja como ejercicio 32.

En la sección 7.4 se descubrirá una fórmula para el operador norma $\|A\|_2$, aunque no es tan factible computacionalmente como la fórmula para $\|A\|_1$ o $\|A\|_\infty$.

El número de condición de una matriz

En la Exploración: "Mentiras que me dice mi computadora" del capítulo 2, se encontró la noción de un sistema *mal condicionado* de ecuaciones lineales. He aquí la definición como se aplica a matrices.

Definición Una matriz A está ***mal condicionada*** si pequeños cambios en sus entradas pueden producir grandes cambios en las soluciones de $A\mathbf{x} = \mathbf{b}$. Si pequeños cambios en las entradas de A sólo producen pequeños cambios en las soluciones de $A\mathbf{x} = \mathbf{b}$, entonces A se llama ***bien condicionada***.

Aunque la definición se aplica a matrices arbitrarias, la atención se centrará en las matrices cuadradas.

Ejemplo 7.20

Demuestre que $A = \begin{bmatrix} 1 & 1 \\ 1 & 1.0005 \end{bmatrix}$ está mal condicionada.

Solución Si toma $\mathbf{b} = \begin{bmatrix} 3 \\ 3.0010 \end{bmatrix}$, entonces la solución a $A\mathbf{x} = \mathbf{b}$ es $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Sin embargo, si A cambia a

$$A' = \begin{bmatrix} 1 & 1 \\ 1 & 1.0010 \end{bmatrix}$$

entonces la solución cambia a $\mathbf{x}' = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. (Compruebe estas afirmaciones.) Por tanto, un cambio relativo de $0.0005/1.0005 \approx 0.0005$, o aproximadamente 0.05%, causa un cambio de $(2-1)/1 = 1$, o 100%, en x_1 y $(1-2)/2 = -0.5$, o -50%, en x_2. En consecuencia, A está mal condicionada.

Es posible usar normas matriciales para ofrecer una forma más precisa de determinar cuándo una matriz está mal condicionada. Piense en el cambio de A a A' como en un error ΔA que, a su vez, introduce un error $\Delta\mathbf{x}$ en la solución $\mathbf{x}$ de $A\mathbf{x} = \mathbf{b}$. Entonces $A' = A + \Delta A$ y $\mathbf{x}' = \mathbf{x} + \Delta\mathbf{x}$. En el ejemplo 7.20,

$$\Delta A = \begin{bmatrix} 0 & 0 \\ 0 & 0.0005 \end{bmatrix} \quad \text{y} \quad \Delta\mathbf{x} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Entonces, puesto que $A\mathbf{x} = \mathbf{b}$ y $A'\mathbf{x}' = \mathbf{b}$, se tiene $(A + \Delta A)(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{b}$. Al expandir y cancelar $A\mathbf{x} = \mathbf{b}$, se obtiene

$$A(\Delta\mathbf{x}) + (\Delta A)\mathbf{x} + (\Delta A)(\Delta\mathbf{x}) = \mathbf{0} \quad \text{o} \quad A(\Delta\mathbf{x}) = -\Delta A(\mathbf{x} + \Delta\mathbf{x})$$

Dado que se supuso que $A\mathbf{x} = \mathbf{b}$ tiene una solución, A debe ser invertible. Por tanto, la última ecuación puede rescribirse como

$$\Delta\mathbf{x} = -A^{-1}(\Delta A)(\mathbf{x} + \Delta\mathbf{x}) = -A^{-1}(\Delta A)\mathbf{x}'$$

Al aplicar normas en ambos lados (usando una norma matricial que sea compatible con una norma vectorial), se tiene

$$\begin{aligned} \|\Delta\mathbf{x}\| &= \|-A^{-1}(\Delta A)\mathbf{x}'\| = \|A^{-1}(\Delta A)\mathbf{x}'\| \\ &\le \|A^{-1}(\Delta A)\|\|\mathbf{x}'\| \\ &\le \|A^{-1}\|\|\Delta A\|\|\mathbf{x}'\| \end{aligned}$$

(¿Cuál es la justificación para cada paso?) Por tanto,

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}'\|} \le \|A^{-1}\|\|\Delta A\| = (\|A^{-1}\|\|A\|)\frac{\|\Delta A\|}{\|A\|}$$

La expresión $\|A^{-1}\|\|A\|$ se llama ***número de condición*** de A y se denota cond(A). Si A no es invertible, se define cond$(A) = \infty$.

¿Qué se hará con la desigualdad anterior? La razón $\|\Delta A\|/\|A\|$ es una medida del *cambio relativo* en la matriz A, que se supone es pequeño. De igual modo, $\|\Delta\mathbf{x}\|/\|\mathbf{x}'\|$ es una medida del error relativo creado en la solución de $A\mathbf{x} = \mathbf{b}$ (aunque, en este caso, el error se mide en relación con la *nueva* solución, **x**', no con la original, **x**). Por tanto, la desigualdad

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}'\|} \le \text{cond}(A)\frac{\|\Delta A\|}{\|A\|} \tag{1}$$

ofrece una cota superior acerca de cuán grande puede ser el error relativo en la solución en términos del error relativo en la matriz coeficiente. Mientras más grande sea el número de condición, más mal condicionada estará la matriz, pues habrá más "espacio" para que el error sea grande en relación con la solución.

Comentarios

- El número de condición de una matriz depende de la elección de la norma. Las normas más comúnmente usadas son los operadores norma $\|A\|_1$ y $\|A\|_\infty$.
- Para cualquier norma, $\text{cond}(A) \geq 1$. (Vea el ejercicio 45.)

Ejemplo 7.21

Encuentre el número de condición de $A = \begin{bmatrix} 1 & 1 \\ 1 & 1.0005 \end{bmatrix}$ relativo a la ∞-norma.

Solución Calcule primero

$$A^{-1} = \begin{bmatrix} 2001 & -2000 \\ -2000 & 2000 \end{bmatrix}$$

Por tanto, en la ∞-norma (suma absoluta máxima de renglón),

$$\|A\|_\infty = 1 + 1.0005 = 2.0005 \quad \text{y} \quad \|A^{-1}\|_\infty = 2001 + |-2000| = 4001$$

de modo que $\text{cond}_\infty(A) = \|A^{-1}\|_\infty \|A\|_\infty = 4001(2.0005) \approx 8004$.

Es evidente que si el número de condición es grande en relación con una norma matricial compatible, será grande en relación con *cualquier* norma matricial compatible. Por ejemplo, puede demostrarse que para la matriz A de los ejemplos 7.20 y 7.21, $\text{cond}_1(A) \approx 8004$, $\text{cond}_2(A) \approx 8002$ (en relación con la 2-norma) y $\text{cond}_F(A) \approx 8002$ (en relación con la norma de Frobenius).

La convergencia de los métodos iterativos

En la sección 2.5 se exploraron dos métodos iterativos para resolver un sistema de ecuaciones lineales: el método de Jacobi y el método de Gauss-Seidel. En el Teorema 2.9 se enunció sin demostración que si A es una matriz de $n \times n$ estrictamente diagonal dominante, entonces ambos métodos convergen en la solución de $A\mathbf{x} = \mathbf{b}$. Ahora está en posición de demostrar este teorema. De hecho, uno de los usos importantes de las normas matriciales es establecer las propiedades de convergencia de varios métodos iterativos.

Aquí sólo se tratará el método de Jacobi. (El método de Gauss-Seidel puede manejarse usando técnicas similares, pero requiere un poco más de cuidado.) La clave es reescribir el proceso iterativo en términos de matrices. Vuelva a revisar el ejemplo 2.37 con esto en mente. El sistema de ecuaciones es

$$\begin{aligned} 7x_1 - x_2 &= 5 \\ 3x_1 - 5x_2 &= -7 \end{aligned} \tag{2}$$

de modo que $$A = \begin{bmatrix} 7 & -1 \\ 3 & -5 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 5 \\ -7 \end{bmatrix}$$

La ecuación (2) se reescribe como

$$\begin{aligned} x_1 &= \frac{5 + x_2}{7} \\ x_2 &= \frac{7 + 3x_1}{5} \end{aligned} \tag{3}$$

que es equivalente a

$$\begin{aligned} 7x_1 &= x_2 + 5 \\ -5x_2 &= -3x_1 - 7 \end{aligned} \tag{4}$$

o, en términos de matrices,

$$\begin{bmatrix} 7 & 0 \\ 0 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -3 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 5 \\ -7 \end{bmatrix} \tag{5}$$

Estudie cuidadosamente la ecuación (5): la matriz a la izquierda contiene las entradas diagonales de A, mientras que en el lado derecho se ve el *negativo* de las entradas fuera de la diagonal de A y el vector $\mathbf{b}$. De este modo, si A se descompone como

$$A = \begin{bmatrix} 7 & -1 \\ 3 & -5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix} + \begin{bmatrix} 7 & 0 \\ 0 & -5 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} = L + D + U$$

entonces la ecuación (5) puede escribirse como

$$D\mathbf{x} = -(L + U)\mathbf{x} + \mathbf{b}$$

o, de manera equivalente,

$$\mathbf{x} = -D^{-1}(L + U)\mathbf{x} + D^{-1}\mathbf{b} \tag{6}$$

pues la matriz D es invertible. La ecuación (6) es la versión matricial de la ecuación (3). Es fácil ver que es posible hacer esto en general: una matriz A de $n \times n$ puede escribirse como $A = L + D + U$, donde D es la parte diagonal de A, y L y U son, respectivamente, las porciones de A abajo y arriba de la diagonal. El sistema $A\mathbf{x} = \mathbf{b}$ puede escribirse entonces en la forma de la ecuación (6), siempre que D sea invertible, lo que ocurre si A es estrictamente diagonal dominante. (¿Por qué?) Para simplificar la notación, sean $M = -D^{-1}(L + U)$ y $\mathbf{c} = D^{-1}\mathbf{b}$ de modo que la ecuación (6) se convierta en

$$\mathbf{x} = M\mathbf{x} + \mathbf{c} \tag{7}$$

Recuerde cómo usar esta ecuación en el método de Jacobi. Comience con un vector inicial $\mathbf{x}_0$ y colóquelo en el lado derecho de la ecuación (7) para obtener la primera iteración $\mathbf{x}_1$; esto es, $\mathbf{x}_1 = M\mathbf{x}_0 + \mathbf{c}$. Luego coloque $\mathbf{x}_1$ en el lado derecho de la ecuación (7) para obtener la segunda iteración $\mathbf{x}_2 = M\mathbf{x}_1 + \mathbf{c}$. En general, se tiene

$$\mathbf{x}_{k+1} = M\mathbf{x}_k + \mathbf{c} \tag{8}$$

para $k \geq 0$. Por el ejemplo 2.37, se tiene

$$M = -D^{-1}(L + U) = -\begin{bmatrix} 7 & 0 \\ 0 & -5 \end{bmatrix}^{-1}\begin{bmatrix} 0 & -1 \\ 3 & 0 \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{7} \\ \frac{3}{5} & 0 \end{bmatrix}$$

y

$$\mathbf{c} = D^{-1}\mathbf{b} = \begin{bmatrix} 7 & 0 \\ 0 & -5 \end{bmatrix}^{-1}\begin{bmatrix} 5 \\ -7 \end{bmatrix} = \begin{bmatrix} \frac{5}{7} \\ \frac{7}{5} \end{bmatrix}$$

de modo que
$$\mathbf{x}_1 = \begin{bmatrix} 0 & \frac{1}{7} \\ \frac{3}{5} & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{5}{7} \\ \frac{7}{5} \end{bmatrix} = \begin{bmatrix} \frac{5}{7} \\ \frac{7}{5} \end{bmatrix} \approx \begin{bmatrix} 0.714 \\ 1.400 \end{bmatrix}$$
$$\mathbf{x}_2 = \begin{bmatrix} 0 & \frac{1}{7} \\ \frac{3}{5} & 0 \end{bmatrix}\begin{bmatrix} 0.714 \\ 1.400 \end{bmatrix} + \begin{bmatrix} \frac{5}{7} \\ \frac{7}{5} \end{bmatrix} \approx \begin{bmatrix} 0.914 \\ 1.829 \end{bmatrix}$$

etcétera. (Éstos son exactamente los mismos cálculos que se hicieron en el ejemplo 2.37, pero escritos en forma matricial.)

Para demostrar que el método de Jacobi convergirá, es necesario demostrar que las iteraciones $\mathbf{x}_k$ tienden a la solución real $\mathbf{x}$ de $A\mathbf{x} = \mathbf{b}$. Es suficiente con demostrar que los ***vectores de error*** $\mathbf{x}_k - \mathbf{x}$ tienden al vector cero. A partir de los cálculos anteriores, $A\mathbf{x} = \mathbf{b}$ es equivalente a $\mathbf{x} = M\mathbf{x} + \mathbf{c}$. Al usar la ecuación (8), se tiene entonces

$$\begin{aligned} \mathbf{x}_{k+1} - \mathbf{x} &= M\mathbf{x}_k + \mathbf{c} - (M\mathbf{x} + \mathbf{c}) \\ &= M(\mathbf{x}_k - \mathbf{x}) \end{aligned}$$

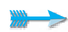
Ahora se aplica la norma en ambos lados de esta ecuación. (En este punto no es importante cuál norma se use en tanto se elija una norma matricial que sea compatible con una norma vectorial.) Se tiene

$$\|\mathbf{x}_{k+1} - \mathbf{x}\| = \|M(\mathbf{x}_k - \mathbf{x})\| \leq \|M\|\|\mathbf{x}_k - \mathbf{x}\| \tag{9}$$

Si puede demostrar que $\|M\| < 1$, entonces tendrá $\|\mathbf{x}_{k+1} - \mathbf{x}\| < \|\mathbf{x}_k - \mathbf{x}\|$ para toda $k \geq 0$, y se tiene que $\|\mathbf{x}_k - \mathbf{x}\|$ tiende a cero, de modo que los vectores de error $\mathbf{x}_k - \mathbf{x}$ tienden al vector cero.

El hecho de que la dominancia diagonal estricta se defina en términos de los valores absolutos de las entradas en los *renglones* de una matriz sugiere que debe elegirse la ∞-norma de una matriz (el operador norma inducido por la norma máx). Si $A = [a_{ij}]$, entonces

$$M = \begin{bmatrix} 0 & -a_{12}/a_{11} & \cdots & -a_{1n}/a_{11} \\ -a_{21}/a_{22} & 0 & \cdots & -a_{2n}/a_{22} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1}/a_{nn} & -a_{n2}/a_{nn} & \cdots & 0 \end{bmatrix}$$



(verifique esto), de modo que, por el Teorema 7.7, $\|M\|_\infty$ es la suma absoluta máxima de renglón de M. Suponga que ello ocurre en el k-ésimo renglón. Entonces

$$\begin{aligned} \|M\|_\infty &= \left|\frac{-a_{k1}}{a_{kk}}\right| + \cdots + \left|\frac{-a_{k,\,k-1}}{a_{kk}}\right| + \left|\frac{-a_{k,\,k+1}}{a_{kk}}\right| + \cdots + \left|\frac{-a_{kn}}{a_{kk}}\right| \\ &= \frac{|a_{k1}| + \cdots + |a_{k,\,k-1}| + |a_{k,\,k+1}| + \cdots + |a_{kn}|}{|a_{kk}|} < 1 \end{aligned}$$

pues A es estrictamente diagonal dominante. Por ende, $\|M\|_\infty < 1$, de modo que $\|\mathbf{x}_k - \mathbf{x}\| \to 0$, como se quería demostrar.

Ejemplo 7.22

Calcule $\|M\|_\infty$ en el ejemplo 2.37 y use este valor para encontrar el número de iteraciones requeridas para aproximar la solución a una precisión de tres lugares decimales (después de redondear) si el vector inicial es $\mathbf{x}_0 = \mathbf{0}$.

Solución Ya se calculó $M = \begin{bmatrix} 0 & \frac{1}{7} \\ \frac{3}{5} & 0 \end{bmatrix}$, de modo que $\|M\|_\infty = \frac{3}{5} = 0.6 < 1$ (lo que implica que el método de Jacobi converge en el ejemplo 2.37, como se vio). La solución aproximada $\mathbf{x}_k$ será precisa a tres lugares decimales si el vector de error $\mathbf{x}_k - \mathbf{x}$ tiene la propiedad de que cada uno de sus componentes es menor que 0.0005 en valor absoluto. (¿Por qué?) Por tanto, sólo es necesario garantizar que el componente absoluto *máximo* de $\mathbf{x}_k - \mathbf{x}$ sea menor que 0.0005. En otras palabras, es necesario encontrar el valor más pequeño de k tal que

$$\|\mathbf{x}_k - \mathbf{x}\|_m < 0.0005$$

Al usar la ecuación (9) anterior se ve que

$$\|\mathbf{x}_k - \mathbf{x}\|_m \leq \|M\|_\infty \|\mathbf{x}_{k-1} - \mathbf{x}\|_m \leq \|M\|_\infty^2 \|\mathbf{x}_{k-2} - \mathbf{x}\|_m \leq \cdots \leq \|M\|_\infty^k \|\mathbf{x}_0 - \mathbf{x}\|_m$$

Ahora $\|M\|_\infty = 0.6$ y $\|\mathbf{x}_0 - \mathbf{x}\|_m \approx \|\mathbf{x}_0 - \mathbf{x}_1\|_m = \|\mathbf{x}_1\|_m = \left\| \begin{bmatrix} 0.714 \\ 1.400 \end{bmatrix} \right\|_m = 1.4$, de modo que

$$\|M\|_\infty^k \|\mathbf{x}_0 - \mathbf{x}\|_m \approx (0.6)^k (1.4)$$

(Si se conoce la solución exacta por adelantado, podría usarse en lugar de $\mathbf{x}_1$. En la práctica, este no es el caso, así que se usa una aproximación de la solución, como se hizo aquí.) Por tanto, es necesario encontrar k tal que

$$(0.6)^k(1.4) < 0.0005$$

Puede resolver esta desigualdad al aplicar logaritmos (base 10) en ambos lados. Se tiene

$$\begin{aligned} \log_{10}((0.6)^k(1.4)) < \log_{10}(5 \times 10^{-4}) &\Rightarrow k\log_{10}(0.6) + \log_{10}(1.4) < \log_{10}5 - 4 \\ &\Rightarrow -0.222k + 0.146 < -3.301 \\ &\Rightarrow k > 15.5 \end{aligned}$$

Puesto que k debe ser entero, en consecuencia se puede concluir que $k = 16$ funcionará y que 16 iteraciones del método de Jacobi darán una precisión a tres lugares decimales en este ejemplo. (De hecho, a partir de los cálculos en el ejemplo 2.37, parece que este grado de precisión se consigue más pronto, pero la meta aquí sólo era encontrar una estimación.)

Ejercicios 7.2

En los ejercicios 1-3, sean $\mathbf{u} = \begin{bmatrix} -1 \\ 4 \\ -5 \end{bmatrix}$ *y* $\mathbf{v} = \begin{bmatrix} 2 \\ -2 \\ 0 \end{bmatrix}$.

1. Calcule la norma euclidiana, la norma suma y la norma máx de $\mathbf{u}$.

2. Calcule la norma euclidiana, la norma suma y la norma máx de $\mathbf{v}$.

3. Calcule $d(\mathbf{u}, \mathbf{v})$ en relación con la norma euclidiana, la norma suma y la norma máx.

4. **(a)** ¿Qué mide $d_s(\mathbf{u}, \mathbf{v})$?
(b) ¿Qué mide $d_m(\mathbf{u}, \mathbf{v})$?

En los ejercicios 5 y 6, sean $\mathbf{u} = [1 \quad 0 \quad 1 \quad 1 \quad 0 \quad 0 \quad 1]^T$ *y* $\mathbf{v} = [0 \quad 1 \quad 1 \quad 0 \quad 1 \quad 1 \quad 1]^T$.

5. Calcule las normas de Hamming de $\mathbf{u}$ y $\mathbf{v}$.

6. Calcule la distancia de Hamming entre $\mathbf{u}$ y $\mathbf{v}$.

7. **(a)** ¿Para cuáles vectores $\mathbf{v}$ es $\|\mathbf{v}\|_E = \|\mathbf{v}\|_m$? Explique su respuesta.

(b) ¿Para cuáles vectores $\mathbf{v}$ es $\|\mathbf{v}\|_s = \|\mathbf{v}\|_m$? Explique su respuesta.

(c) ¿Para cuáles vectores $\mathbf{v}$ es $\|\mathbf{v}\|_s = \|\mathbf{v}\|_m = \|\mathbf{v}\|_E$? Explique su respuesta.

8. (a) ¿En qué condiciones para $\mathbf{u}$ y $\mathbf{v}$ es $\|\mathbf{u} + \mathbf{v}\|_E = \|\mathbf{u}\|_E + \|\mathbf{v}\|_E$? Explique su respuesta.

(b) ¿En qué condiciones para $\mathbf{u}$ y $\mathbf{v}$ es $\|\mathbf{u} + \mathbf{v}\|_s = \|\mathbf{u}\|_s + \|\mathbf{v}\|_s$? Explique su respuesta.

(c) ¿En qué condiciones para $\mathbf{u}$ y $\mathbf{v}$ es $\|\mathbf{u} + \mathbf{v}\|_m = \|\mathbf{u}\|_m + \|\mathbf{v}\|_m$? Explique su respuesta.

9. Demuestre que para todo $\mathbf{v}$ en $\mathbb{R}^n$, $\|\mathbf{v}\|_m \le \|\mathbf{v}\|_E$.

10. Demuestre que para todo $\mathbf{v}$ en $\mathbb{R}^n$, $\|\mathbf{v}\|_E \le \|\mathbf{v}\|_s$.

11. Demuestre que para todo $\mathbf{v}$ en $\mathbb{R}^n$, $\|\mathbf{v}\|_s \le n\|\mathbf{v}\|_m$.

12. Demuestre que para todo $\mathbf{v}$ en $\mathbb{R}^n$, $\|\mathbf{v}\|_E \le \sqrt{n}\|\mathbf{v}\|_m$.

13. Dibuje los círculos unitarios en $\mathbb{R}^2$ en relación con la norma suma y la norma máx.

14. Al demostrar que la identidad del ejercicio 33 en la sección 7.1 fracasa, demuestre que la norma suma no surge de cualquier producto interno.

En los ejercicios 15-18, pruebe que $\|\ \|$ define una norma sobre el espacio vectorial V.

15. $V = \mathbb{R}^2, \left\|\begin{bmatrix} a \\ b \end{bmatrix}\right\| = \text{máx}\{|2a|, |3b|\}$

16. $V = M_{mn}, \|A\| = \max\limits_{i,j}\{|a_{ij}|\}$

17. $V = \mathscr{C}[0, 1], \|f\| = \int_0^1 |f(x)|\,dx$

18. $\|f\| = \max\limits_{0\le x\le 1} |f(x)|$

19. Demuestre el Teorema 7.5(b).

En los ejercicios 20-25, calcule $\|A\|_F$, $\|A\|_1$ y $\|A\|_\infty$.

20. $A = \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}$

21. $A = \begin{bmatrix} 0 & -1 \\ -3 & 3 \end{bmatrix}$

22. $A = \begin{bmatrix} 1 & 5 \\ -2 & -1 \end{bmatrix}$

23. $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 1 & 3 \end{bmatrix}$

24. $A = \begin{bmatrix} 0 & -5 & 2 \\ 3 & 1 & -3 \\ -4 & -4 & 3 \end{bmatrix}$

25. $A = \begin{bmatrix} 4 & -2 & -1 \\ 0 & -1 & 2 \\ 3 & -3 & 0 \end{bmatrix}$

En los ejercicios 26-31, encuentre vectores $\mathbf{x}$ *y* $\mathbf{y}$ *con* $\|\mathbf{x}\|_s = 1$ *y* $\|\mathbf{y}\|_m = 1$ *tales que* $\|A\|_1 = \|A\mathbf{x}\|_s$ *y* $\|A\|_\infty = \|A\mathbf{y}\|_m$, *donde A es la matriz en el ejercicio dado.*

26. Ejercicio 20 **27.** Ejercicio 21 **28.** Ejercicio 22

29. Ejercicio 23 **30.** Ejercicio 24 **31.** Ejercicio 25

32. Demuestre el Teorema 7.7(b).

33. (a) Si $\|A\|$ es un operador norma, pruebe que $\|I\| = 1$, donde I es una matriz identidad.

(b) ¿Existe alguna norma vectorial que induzca la norma de Frobenius como un operador norma? ¿Por qué sí o por qué no?

34. Sea $\|A\|$ una norma matricial que sea compatible con una norma vectorial $\|\mathbf{x}\|$. Pruebe que $\|A\| \ge |\lambda|$ para todo eigenvalor λ de A.

En los ejercicios 35-40, encuentre $\text{cond}_1(A)$ *y* $\text{cond}_\infty(A)$. *Establezca si la matriz dada está mal condicionada.*

35. $A = \begin{bmatrix} 3 & 1 \\ 4 & 2 \end{bmatrix}$

36. $A = \begin{bmatrix} 1 & -2 \\ -3 & 6 \end{bmatrix}$

37. $A = \begin{bmatrix} 1 & 0.99 \\ 1 & 1 \end{bmatrix}$

38. $A = \begin{bmatrix} 150 & 200 \\ 3001 & 4002 \end{bmatrix}$

39. $A = \begin{bmatrix} 1 & 1 & 1 \\ 5 & 5 & 6 \\ 1 & 0 & 0 \end{bmatrix}$

40. $A = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{bmatrix}$

41. Sea $A = \begin{bmatrix} 1 & k \\ 1 & 1 \end{bmatrix}$.

(a) Encuentre una fórmula para $\text{cond}_\infty(A)$ en términos de k.

(b) ¿Qué le ocurre a $\text{cond}_\infty(A)$ conforme k tiende a 1?

42. Considere el sistema lineal $A\mathbf{x} = \mathbf{b}$, donde A es invertible. Suponga que un error $\Delta\mathbf{b}$ cambia $\mathbf{b}$ a $\mathbf{b}' = \mathbf{b} + \Delta\mathbf{b}$ Sea $\mathbf{x}'$ la solución del nuevo sistema; esto es, $A\mathbf{x}' = \mathbf{b}'$. Sea $\mathbf{x}' = \mathbf{x} + \Delta\mathbf{x}$ de modo que $\Delta\mathbf{x}$ representa el error resultante en la solución del sistema. Demuestre que

$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} \le \text{cond}(A)\frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}$$

para cualquier norma matricial compatible.

43. Sean $A = \begin{bmatrix} 10 & 10 \\ 10 & 9 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 100 \\ 99 \end{bmatrix}$.

(a) Calcule $\text{cond}_\infty(A)$.

(b) Suponga que A cambia a $A' = \begin{bmatrix} 10 & 10 \\ 10 & 11 \end{bmatrix}$. ¿Cuán grande es el cambio relativo que puede producir este cambio en la solución de $A\mathbf{x} = \mathbf{b}$? [*Sugerencia:* use la desigualdad (1) de esta sección.]

(c) Resuelva los sistemas usando A y A' y determine el error relativo real.

(d) Suponga que $\mathbf{b}$ cambia a $\mathbf{b}' = \begin{bmatrix} 100 \\ 101 \end{bmatrix}$. ¿Cuán grande es el cambio relativo que puede producir este cambio en la solución de $A\mathbf{x} = \mathbf{b}$? [*Sugerencia:* use el ejercicio 42.]

(e) Resuelva los sistemas usando $\mathbf{b}$ y $\mathbf{b}'$ y determine el error relativo real.

44. Sean $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 5 & 0 \\ 1 & -1 & 2 \end{bmatrix}$ y $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

(a) Calcule $\text{cond}_1(A)$.

(b) Suponga que A cambia a $A' = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 5 & 0 \\ 1 & -1 & 2 \end{bmatrix}$.

¿Cuán grande es el cambio relativo que puede producir este cambio en la solución de $A\mathbf{x} = \mathbf{b}$?[*Sugerencia:* use la desigualdad (1) de esta sección.]

(c) Resuelva los sistemas usando A y A' y determine el error relativo real.

(d) Suponga que $\mathbf{b}$ cambia a $\mathbf{b}' = \begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}$. ¿Cuán grande es el cambio relativo que puede producir este cambio en la solución de $A\mathbf{x} = \mathbf{b}$? [*Sugerencia:* use el ejercicio 42.]

(e) Resuelva los sistemas usando $\mathbf{b}$ y $\mathbf{b}'$ y determine el error relativo real.

45. Demuestre que si A es una matriz invertible, entonces $\text{cond}(A) \geq 1$ con respecto a cualquier norma matricial.

46. Demuestre que si A y B son matrices invertibles, entonces $\text{cond}(AB) \leq \text{cond}(A)\,\text{cond}(B)$ con respecto a cualquier norma matricial.

47. Sea A una matriz invertible y sean λ_1 y λ_2 los eigenvalores con los valores absolutos más grande y más pequeño, respectivamente. Demuestre que

$$\text{cond}\,(A) \geq \frac{|\lambda_1|}{|\lambda_n|}$$

[*Sugerencia:* vea el ejercicio 34 y el Teorema 4.18(b) en la sección 4.3.]

CAS *En los ejercicios 48-51, escriba el sistema dado en la forma de la ecuación (7). Luego use el método del ejemplo 7.22 para estimar el número de iteraciones del método de Jacobi que se necesitarán para aproximar la solución a una precisión de tres lugares decimales. (Use $\mathbf{x}_0 = \mathbf{0}$.) Compare su respuesta con la solución calculada en el ejercicio dado de la sección 2.5.*

48. Ejercicio 1, sección 2.5
49. Ejercicio 3, sección 2.5
50. Ejercicio 4, sección 2.5
51. Ejercicio 5, sección 2.5

El ejercicio 52(c) se refiere al modelo de Leontief de una economía abierta, que se estudió en las secciones 2.4 y 3.7.

52. Sea A una matriz de $n \times n$ tal que $\|A\| < 1$, donde la norma es la norma suma o la norma máx.

(a) Pruebe que $A^n \to O$ conforme $n \to \infty$.

(b) Deduzca a partir de (a) que $I - A$ es invertible y

$$(I - A)^{-1} = I + A + A^2 + A^3 + \cdots$$

[*Sugerencia:* Vea la demostración del Teorema 3.34.]

(c) Demuestre que (b) puede usarse para probar los Corolarios 3.35 y 3.36.

7.3 Aproximación por mínimos cuadrados

En muchas ramas de la ciencia, los datos experimentales se usan para inferir una relación matemática entre las variables por medir. Por ejemplo, puede medir la altura de un árbol en varios puntos en el tiempo y tratar de deducir una función que exprese la altura h del árbol en términos del tiempo t. O puede medir el tamaño p de una población a lo largo de tiempo y tratar de encontrar una regla que relacione p con t. Las relaciones entre variables también son de interés en los negocios; por ejemplo, una compañía que produce algunos objetos puede estar interesada en conocer la relación entre sus costos totales c y el número n de objetos producidos.

En cada uno de estos ejemplos, los datos vienen en forma de dos mediciones: una para la variable independiente y uno para la (supuesta) variable dependiente. Por ende, se tiene un conjunto de *puntos de datos* (x_i, y_i) y se busca una función que aproxime mejor la relación entre la variable independiente x y la variable dependiente y. La figura 7.9 muestra ejemplos en los que los puntos de datos experimentales se grafican, junto con una curva que "ajusta" aproximadamente los datos.

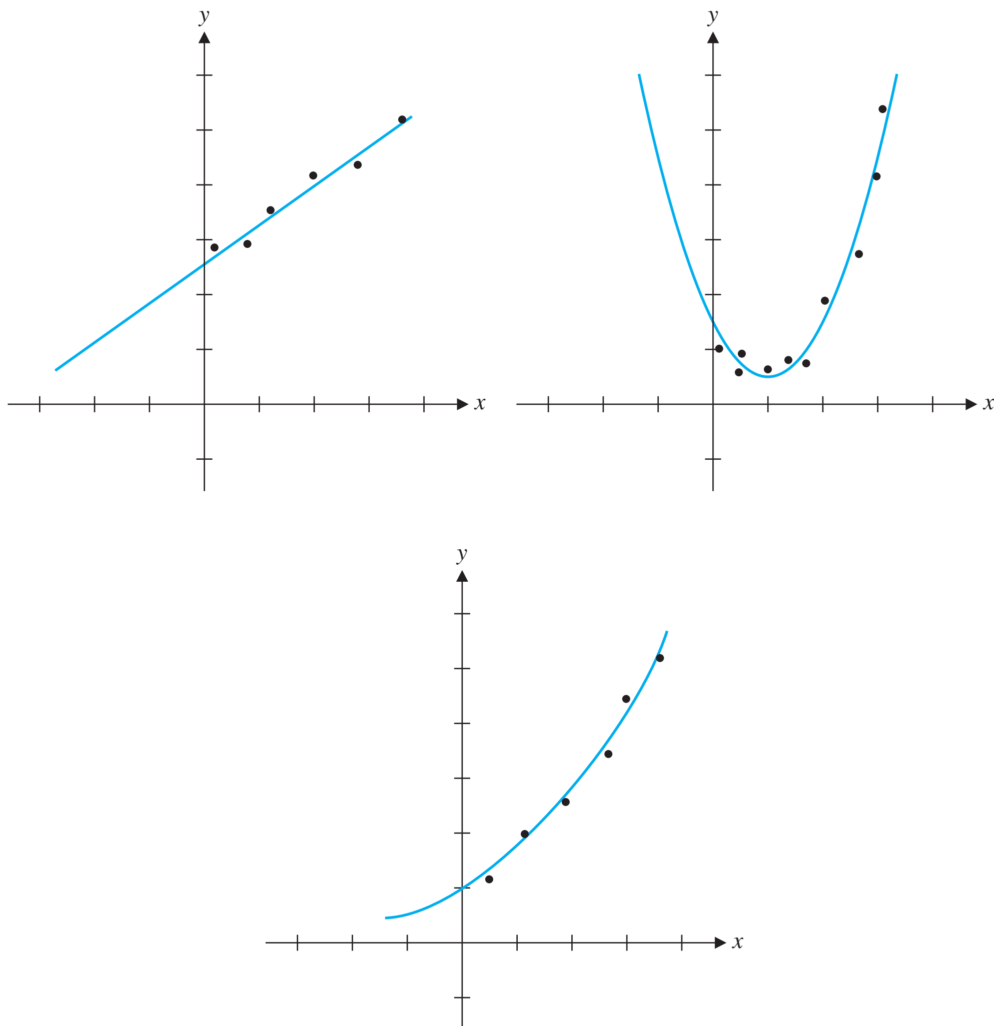


Figura 7.9
Curvas de "mejor ajuste"

Roger Cotes (1682-1716) fue un matemático inglés quien, mientras impartía cátedra en Cambridge, editó la segunda edición de los *Principia* de Newton. Aunque publicó poco, hizo importantes descubrimientos en la teoría de logaritmos, cálculo integral y métodos numéricos.

El método de mínimos cuadrados, que está a punto de considerarse, se atribuye a Gauss. En el día de año nuevo de 1801 se descubrió un nuevo asteroide, Ceres, pero desapareció detrás del Sol poco después de observarse. Los astrónomos predijeron cuándo y dónde reaparecería Ceres, pero sus cálculos difirieron enormemente de los realizados, de manera independiente, por Gauss. Ceres reapareció el 7 de diciembre de 1801, casi exactamente donde Gauss predijo que lo haría. Aunque él no divulgó su método en el momento, usó su método de aproximación por mínimos cuadrados, que describió en un artículo en 1809. En realidad, el mismo método se conocía anteriormente; Cotes anticipó el método a principios del siglo XVIII, y Legendre publicó un artículo acerca del mismo en 1806. No obstante, por lo general se acredita a Gauss el método de aproximación por mínimos cuadrados.

La exploración de la aproximación comienza con un resultado más general.

El teorema de mejor aproximación

En ciencias existen muchos problemas que pueden enunciarse generalmente como “¿cuál es la mejor aproximación a X del tipo Y?” X puede ser un conjunto de puntos de datos, una función, un vector o muchas otras cosas, mientras que Y puede ser un tipo particular de función, un vector que pertenezca a cierto espacio vectorial, etcétera. Un ejemplo típico de tal problema es encontrar el vector **w** en un subespacio W de un espacio vectorial V que se aproxime mejor (es decir, esté más cerca)a un vector dado **v** en V. Este problema da lugar a la siguiente definición.

Definición Si W es un subespacio de un espacio lineal normado V y si **v** es un vector en V, entonces la ***mejor aproximación a* v *en* W** es el vector $\overline{\mathbf{v}}$ en W tal que

$$\|\mathbf{v} - \overline{\mathbf{v}}\| < \|\mathbf{v} - \mathbf{w}\|$$

para todo vector **w** en W diferente de $\overline{\mathbf{v}}$.

En $\mathbb{R}^2$ o $\mathbb{R}^3$, uno está acostumbrado a pensar en “distancia más corta” como lo correspondiente a “distancia perpendicular”. En terminología algebraica, “distancia más corta” se relaciona con la noción de proyección ortogonal: si W es un subespacio de $\mathbb{R}^n$ y **v** es un vector en $\mathbb{R}^n$, entonces se espera que $\text{proy}_W(\mathbf{v})$ sea el vector en W que esté más cerca de **v** (figura 7.10).

Puesto que la proyección ortogonal puede definirse en cualquier espacio con producto interno, se tiene el siguiente teorema.

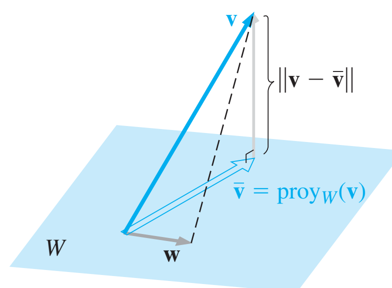


Figura 7.10
Si $\overline{\mathbf{v}} = \text{proy}_W(\mathbf{v})$, entonces
$\|\mathbf{v} - \overline{\mathbf{v}}\| < \|\mathbf{v} - \mathbf{w}\|$ para todo $\mathbf{w} \neq \overline{\mathbf{v}}$

Teorema 7.8 **El teorema de mejor aproximación**

Si W es un subespacio con dimensión finita de un espacio con producto interno V, y si **v** s un vector en V, entonces $\text{proy}_W(\mathbf{v})$ es la mejor aproximación a **v** en W.

Demostración Sea **w** un vector en W diferente de $\text{proy}_W(\mathbf{v})$. Entonces $\text{proy}_W(\mathbf{v}) - \mathbf{w}$ también está en W, de modo que $\mathbf{v} - \text{proy}_W(\mathbf{v}) = \text{perp}_W(\mathbf{v})$ es ortogonal a $\text{proy}_W(\mathbf{v}) - \mathbf{w}$, por el ejercicio 43 de la sección 7.1. Ahora el teorema de Pitágoras implica que

$$\begin{aligned}\|\mathbf{v} - \text{proy}_W(\mathbf{v})\|^2 + \|\text{proy}_W(\mathbf{v}) - \mathbf{w}\|^2 &= \|(\mathbf{v} - \text{proy}_W(\mathbf{v})) + (\text{proy}_W(\mathbf{v}) - \mathbf{w})\|^2 \\ &= \|\mathbf{v} - \mathbf{w}\|^2\end{aligned}$$

como ilustra la figura 7.10. Sin embargo, $\|\text{proy}_W(\mathbf{v}) - \mathbf{w}\|^2 > 0$, pues $\mathbf{w} \neq \text{proy}_W(\mathbf{v})$, de modo que

$$\|\mathbf{v} - \text{proy}_W(\mathbf{v})\|^2 < \|\mathbf{v} - \text{proy}_W(\mathbf{v})\|^2 + \|\text{proy}_W(\mathbf{v}) - \mathbf{w}\|^2 = \|\mathbf{v} - \mathbf{w}\|^2$$

o, de manera equivalente,

$$\|\mathbf{v} - \text{proy}_W(\mathbf{v})\| < \|\mathbf{v} - \mathbf{w}\|$$

Ejemplo 7.23

Sean $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$. Encuentre la mejor aproximación para $\mathbf{v}$ en el plano $W = \text{gen}(\mathbf{u}_1, \mathbf{u}_2)$ y encuentre la distancia euclidiana de $\mathbf{v}$ a W.

Solución El vector en W que aproxima mejor a $\mathbf{v}$ es $\text{proy}_W(\mathbf{v})$. Puesto que $\mathbf{u}_1$ y $\mathbf{u}_2$ son ortogonales,

$$\begin{aligned} \text{proy}_W(\mathbf{v}) &= \left(\frac{\mathbf{u}_1 \cdot \mathbf{v}}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right)\mathbf{u}_1 + \left(\frac{\mathbf{u}_2 \cdot \mathbf{v}}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right)\mathbf{u}_2 \\ &= \tfrac{2}{6}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \tfrac{16}{30}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -\frac{2}{5} \\ \frac{1}{5} \end{bmatrix} \end{aligned}$$

La distancia de $\mathbf{v}$ a W es la distancia de $\mathbf{v}$ al punto en W más cercano a $\mathbf{v}$. Pero esta distancia es justo $\|\text{perp}_W(\mathbf{v})\| = \|\mathbf{v} - \text{proy}_W(\mathbf{v})\|$. Calcule

$$\mathbf{v} - \text{proy}_W(\mathbf{v}) = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} - \begin{bmatrix} 3 \\ -\frac{2}{5} \\ \frac{1}{5} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{12}{5} \\ \frac{24}{5} \end{bmatrix}$$

de modo que $\|\mathbf{v} - \text{proy}_W(\mathbf{v})\| = \sqrt{0^2 + (\frac{12}{5})^2 + (\frac{24}{5})^2} = \sqrt{\frac{720}{25}} = 12\sqrt{5}/5$

que es la distancia de $\mathbf{v}$ a W.

En la sección 7.5 se observarán otros ejemplos del teorema de mejor aproximación cuando se explore el problema de aproximar funciones.

Comentario La proyección ortogonal de un vector $\mathbf{v}$ sobre un subespacio W se define en términos de una base ortogonal para W. El teorema de mejor aproximación proporciona una prueba alternativa de que $\text{proy}_W(\mathbf{v})$ no depende de la elección de esta base, pues puede haber sólo un vector en W esté más cercano a $\mathbf{v}$, a saber, $\text{proy}_W(\mathbf{v})$.

Aproximación por mínimos cuadrados

Ahora regrese al problema de encontrar una curva que "ajuste mejor" un conjunto de puntos de datos. Sin embargo, antes de proceder, es necesario definir qué se entiende por "mejor ajuste". Suponga que los puntos de datos $(1, 2)$, $(2, 2)$ y $(3, 4)$ surgieron de mediciones tomadas durante algún experimento. Suponga también que existen razones para creer que los valores x y y se relacionan mediante un función lineal; esto es, se espera que los puntos se encuentren en alguna recta con ecuación $y = a + bx$. Si las mediciones son precisas, los tres puntos satisfarán esta ecuación y tendrá

$$2 = a + b \cdot 1 \qquad 2 = a + b \cdot 2 \qquad 4 = a + b \cdot 3$$

Éste es un sistema de ecuaciones lineales con dos variables:

$$\begin{aligned} a + b &= 2 \\ a + 2b &= 2 \\ a + 3b &= 4 \end{aligned} \quad \text{o} \quad \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

Desafortunadamente, este sistema es inconsistente (pues los tres puntos no yacen sobre una recta). De modo que podrá conformarse con una recta que pase "tan cerca como sea posible" por los puntos. Para cualquier recta, se medirá la distancia vertical desde cada punto de datos a la recta (lo que representa los *errores* en la dirección y) y luego se tratará de elegir la recta que minimice el *error total*. La figura 7.11 ilustra la situación.

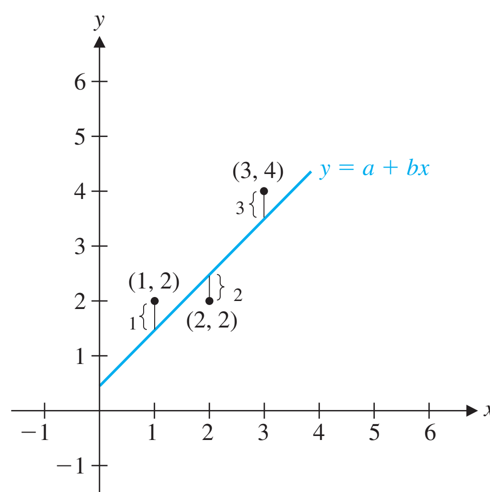


Figura 7.11
Cómo encontrar la recta que minimiza $\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2$

Si los errores se denotan ε_1, ε_2 y ε_3 entonces puede formar el ***vector de error***

$$\mathbf{e} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$$

Se quiere que $\mathbf{e}$ sea tan pequeño como sea posible, de modo que $\|\mathbf{e}\|$ debe estar tan cerca de cero como sea posible. ¿Cuál norma usaría? Es evidente que la familiar norma euclidiana es la mejor elección. (La norma suma también sería una elección sensible, pues $\|\mathbf{e}\|_s = |\varepsilon_1| + |\varepsilon_2| + |\varepsilon_3|$ es la suma real de los errores en la figura 7.11. Sin embargo, es difícil trabajar con los signos de valor absoluto y, como verá dentro de poco, la elección de la norma euclidiana conduce a algunas fórmulas muy amigables.) Así que se minimizará

$$\|\mathbf{e}\| = \sqrt{\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2} \quad \text{o, de manera equivalente} \quad \|\mathbf{e}\|^2 = \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2$$

De aquí proviene el término "mínimos cuadrados": es necesario encontrar la menor suma de cuadrados, en el sentido de la ecuación anterior. El número $\|\mathbf{e}\|$ se llama ***error de mínimos cuadrados*** de la aproximación.

A partir de la figura 7.11 también se obtienen las siguientes fórmulas para ε_1, ε_2 y ε_3 en el ejemplo:

$$\varepsilon_1 = 2 - (a + b\cdot 1) \quad \varepsilon_2 = 2 - (a + b\cdot 2) \quad \varepsilon_3 = 4 - (a + b\cdot 3)$$

Ejemplo 7.24

¿Cuál de las siguientes rectas ofrece el menor error de mínimos cuadrados para los puntos de datos (1, 2), (2, 2) y (3, 4)?

(a) $y = 1 + x$

(b) $y = -2 + 2x$

(c) $y = \frac{2}{3} + x$

Solución La tabla 7.1 muestra los cálculos necesarios.

Tabla 7.1

	$y = 1 + x$	$y = -2 + 2x$	$y = \frac{2}{3} + x$
ε_1	$2 - (1 + 1) = 0$	$2 - (-2 + 2) = 2$	$2 - (\frac{2}{3} + 1) = \frac{1}{3}$
ε_2	$2 - (1 + 2) = -1$	$2 - (-2 + 4) = 0$	$2 - (\frac{2}{3} + 2) = -\frac{2}{3}$
ε_3	$4 - (1 + 3) = 0$	$4 - (-2 + 6) = 0$	$4 - (\frac{2}{3} + 3) = \frac{1}{3}$
$\varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2$	$0^2 + (-1)^2 + 0^2 = 1$	$2^2 + 0^2 + 0^2 = 4$	$(\frac{1}{3})^2 + (-\frac{2}{3})^2 + (\frac{1}{3})^2 = \frac{2}{3}$
$\|\mathbf{e}\|$	1	2	$\sqrt{\frac{2}{3}} \approx 0.816$

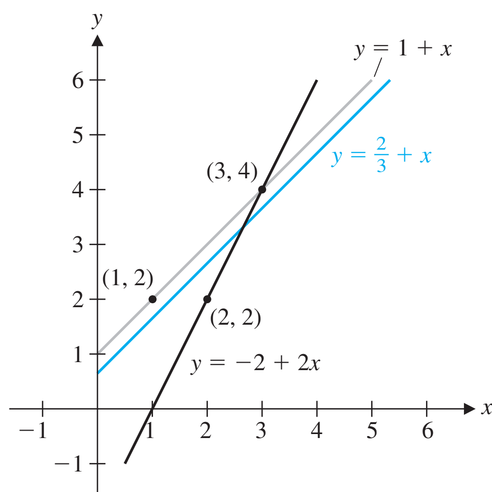


Figura 7.12

Se ve que la recta $y = \frac{2}{3} + x$ produce el menor error de mínimos cuadrados entre estas tres rectas. La figura 7.12 muestra los puntos de datos y las tres rectas.

Es evidente que la recta $y = \frac{2}{3} + x$ del ejemplo 7.24 produce el menor error de mínimos cuadrados que *cualquier* recta, aun cuando *no* pase por ninguno de los puntos dados. El resto de esta sección se dedica a ilustrar por qué esto es así.

En general, suponga que se tienen n puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$ y una recta $y = a + bx$. El vector de error es

$$\mathbf{e} = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

donde $\varepsilon_i = y_i - (a + bx_i)$. La recta $y = a + bx$ que minimiza $\varepsilon_1^2 + \cdots + \varepsilon_n^2$ se llama ***aproximación lineal por mínimos cuadrados*** (o ***recta de mejor ajuste***) para los puntos $(x_1, y_1), \ldots, (x_n, y_n)$. Como se anotó antes del ejemplo 7.24, es posible expresar este problema en forma matricial. Si los puntos dados en realidad están sobre la recta v entonces las n ecuaciones lineales

$$\begin{aligned} a + bx_1 &= y_1 \\ &\vdots \\ a + bx_n &= y_n \end{aligned}$$

serían todas verdaderas (es decir, el sistema sería consistente). El interés está en el caso donde los puntos *no* son colineales, en cuyo caso el sistema es *inconsistente*. En forma matricial se tiene

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

que es de la forma $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$



El vector de error $\mathbf{e}$ es justo $\mathbf{b} - A\mathbf{x}$ (compruébelo) y se quiere minimizar $\|\mathbf{e}\|^2$ o, de manera equivalente, $\|\mathbf{e}\|$. Por tanto, es posible parafrasear el problema en términos de matrices del modo siguiente.

Definición Si A es una matriz de $m \times n$ y $\mathbf{b}$ está en $\mathbb{R}^m$, una ***solución de mínimos cuadrados*** de $A\mathbf{x} = \mathbf{b}$ es un vector $\overline{\mathbf{x}}$ en $\mathbb{R}^n$ tal que

$$\|\mathbf{b} - A\overline{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|$$

para todo $\mathbf{x}$ en $\mathbb{R}^n$.

Solución al problema de mínimos cuadrados

Cualquier vector de la forma $A\mathbf{x}$ está en el espacio columna de A, y conforme $\mathbf{x}$ varía sobre todos los vectores en $\mathbb{R}^n$, $A\mathbf{x}$ varía sobre todos los vectores en col(A). Por tanto, una solución en mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es equivalente a un vector $\overline{\mathbf{y}}$ en col(A) tal que

$$\|\mathbf{b} - \overline{\mathbf{y}}\| \leq \|\mathbf{b} - \mathbf{y}\|$$

para todo $\mathbf{y}$ en col(A). En otras palabras, se necesita el vector más cercano a $\mathbf{b}$ en col(A). Por el teorema de mejor aproximación, el vector que se quiere es la proyección ortogonal de $\mathbf{b}$ sobre col(A). Por ende, si $\overline{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, se tiene

$$A\overline{\mathbf{x}} = \text{proy}_{\text{col}(A)}(\mathbf{b}) \tag{1}$$

Con la finalidad de encontrar $\overline{\mathbf{x}}$, parecería que es necesario primero calcular $\text{proy}_{\text{col}(A)}(\mathbf{b})$ y luego resolver el sistema (1). Sin embargo, hay una mejor forma de proceder.

Se sabe que

$$\mathbf{b} - A\overline{\mathbf{x}} = \mathbf{b} - \text{proy}_{\text{col}(A)}\ \mathbf{b}\ \ = \text{perp}_{\text{col}(A)}\ \mathbf{b})$$

es ortogonal a col(A). De modo que, si $\mathbf{a}_i$ es una columna de A, se tiene

$$\mathbf{a}_i^T(\mathbf{b} - A\overline{\mathbf{x}}) = \mathbf{a}_i \cdot (\mathbf{b} - A\overline{\mathbf{x}}) = 0$$

Esto es cierto si y sólo si

$$\begin{aligned} A^T(\mathbf{b} - A\overline{\mathbf{x}}) &= [\mathbf{a}_1 \quad \cdots \quad \mathbf{a}_n]^T(\mathbf{b} - A\overline{\mathbf{x}}) \\ &= \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_n^T \end{bmatrix}(\mathbf{b} - A\overline{\mathbf{x}}) = \begin{bmatrix} \mathbf{a}_1^T(\mathbf{b} - A\overline{\mathbf{x}}) \\ \vdots \\ \mathbf{a}_n^T(\mathbf{b} - A\overline{\mathbf{x}}) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{0} \end{aligned}$$

que, a su vez, es equivalente a

$$A^T\mathbf{b} - A^TA\overline{\mathbf{x}} = \mathbf{0}$$

o

$$A^TA\overline{\mathbf{x}} = A^T\mathbf{b}$$

Esto representa un sistema de ecuaciones conocido como las ***ecuaciones normales*** para $\overline{\mathbf{x}}$.

Se acaba de establecer que las soluciones de las ecuaciones normales para $\overline{\mathbf{x}}$ son precisamente las soluciones de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$. Esto demuestra la primera parte del siguiente teorema.

Teorema 7.9 **El teorema de mínimos cuadrados**

Sea A una matriz de $m \times n$ y sea $\mathbf{b}$ que está en $\mathbb{R}^m$. Entonces $A\mathbf{x} = \mathbf{b}$ siempre tiene al menos una solución de mínimos cuadrados $\overline{\mathbf{x}}$. Más aún:

a. $\overline{\mathbf{x}}$ es una solución en mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ si y sólo si $\overline{\mathbf{x}}$ es una solución de las ecuaciones normales $A^TA\overline{\mathbf{x}} = A^T\mathbf{b}$.
b. A tiene columnas linealmente independientes si y sólo si A^TA es invertible. En este caso, la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es única y está dada por

$$\overline{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$$

Demostración Ya se estableció la propiedad (a). Para la propiedad (b), note que las n columnas de A son linealmente independientes si y sólo si rank$(A) = n$. Pero esto es verdadero si y sólo si A^TA es invertible, por el Teorema 3.28. Si A^TA es invertible, entonces la única solución de $A^TA\overline{\mathbf{x}} = A^T\mathbf{b}$ claramente es $\overline{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$.

Ejemplo 7.25

Encuentre una solución de mínimos cuadrados para el sistema inconsistente $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}$$

Solución Calcule

$$A^TA = \begin{bmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 0 \\ 0 & 30 \end{bmatrix}$$

y

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 2 \\ 16 \end{bmatrix}$$

Las ecuaciones normales $A^TA\overline{\mathbf{x}} = A^T\mathbf{b}$ son justo

$$\begin{bmatrix} 6 & 0 \\ 0 & 30 \end{bmatrix}\overline{\mathbf{x}} = \begin{bmatrix} 2 \\ 16 \end{bmatrix}$$

que producen $\overline{\mathbf{x}} = \begin{bmatrix} \frac{1}{3} \\ \frac{8}{15} \end{bmatrix}$. El hecho de que esta solución sea única estuvo garantizado por el Teorema 7.9(b), pues las columnas de A claramente son linealmente independientes.

Comentario El ejemplo 7.25 podría parafrasearse del modo siguiente: encuentre la mejor aproximación a $\mathbf{b}$ en el espacio columna de A. Las ecuaciones resultantes proporcionan el sistema $A\mathbf{x} = \mathbf{b}$ cuya solución de mínimos cuadrados acaba de encontrar. (Verifique esto.) En este caso, los componentes de $\overline{\mathbf{x}}$ son los *coeficientes* de aquella combinación lineal de las columnas de A que produce la mejor aproximación a $\mathbf{b}$, a saber,

$$\frac{1}{3}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} + \frac{8}{15}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ -\frac{2}{5} \\ \frac{1}{5} \end{bmatrix}$$

Éste es exactamente el resultado del ejemplo 7.23. Compare los dos planteamientos.

Ejemplo 7.26

Encuentre la aproximación lineal por mínimos cuadrados para los puntos de datos (1, 2), (2, 2) y (3, 4) del ejemplo 7.24.

Solución Ya se vio que el correspondiente sistema $A\mathbf{x} = \mathbf{b}$ es

$$\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix}$$

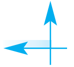
donde $y = a + bx$ es la recta que se busca. Puesto que las columnas de A claramente son linealmente independientes, habrá una solución única de mínimos cuadrados, por el inciso (b) del teorema de mínimos cuadrados. Calcule

$$A^TA = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 6 & 14 \end{bmatrix} \quad \text{y} \quad A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 8 \\ 18 \end{bmatrix}$$

Por tanto, puede resolver las ecuaciones normales $A^TA\overline{\mathbf{x}} = A^T\mathbf{b}$ al usar la eliminación gaussiana para obtener

$$[A^TA \mid A^T\mathbf{b}] = \left[\begin{array}{cc|c} 3 & 6 & 8 \\ 6 & 14 & 18 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 0 & \frac{2}{3} \\ 0 & 1 & 1 \end{array}\right]$$

De modo que $\overline{\mathbf{x}} = \begin{bmatrix} \frac{2}{3} \\ 1 \end{bmatrix}$, de donde se ve que $a = \frac{2}{3}$, $b = 1$ son los coeficientes de la aproximación lineal por mínimos cuadrados: $y = \frac{2}{3} + x$.



La recta recién encontrada es la del ejemplo 7.24(c), así que se justifica la afirmación de que ésta produce el menor error de mínimos cuadrados para los puntos de datos (1, 2), (2, 2) y (3, 4). Note que si $\overline{\mathbf{x}}$ es una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, puede calcular el error de mínimos cuadrados como

$$\|\mathbf{e}\| = \|\mathbf{b} - A\overline{\mathbf{x}}\|$$

Dado que $A\overline{\mathbf{x}} = \text{proy}_{\text{col}(A)}(\mathbf{b})$, esta es justo la longitud de $\text{perp}_{\text{col}(A)}(\mathbf{b})$; es decir, la distancia desde $\mathbf{b}$ hasta el espacio columna de A. En el ejemplo 7.26 se tuvo

$$\mathbf{e} = \mathbf{b} - A\overline{\mathbf{x}} = \begin{bmatrix} 2 \\ 2 \\ 4 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} \frac{2}{3} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

de modo que, como en el ejemplo 7.24(c), se tiene un error de mínimos cuadrados de $\|\mathbf{e}\| = \sqrt{\frac{2}{3}} \approx 0.816$.

Comentario Note que las columnas de A en el ejemplo 7.26 son linealmente independientes, de modo que existe $(A^TA)^{-1}$, y podría calcular $\overline{\mathbf{x}}$ como $\overline{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$. Sin embargo, casi siempre es más sencillo resolver las ecuaciones normales usando eliminación gaussiana (¡o deje que su CAS lo haga por usted!)

Es interesante observar el ejemplo 7.26 desde dos diferentes puntos de vista geométricos. Por un lado, se tiene la aproximación lineal por mínimos cuadrados $y = \frac{2}{3} + x$, con errores correspondientes $\varepsilon_1 = \frac{1}{3}$, $\varepsilon_2 = -\frac{2}{3}$ y $\varepsilon_3 = \frac{1}{3}$, como se muestra en la figura 7.13(a). De manera equivalente, se tiene la proyección de $\mathbf{b}$ sobre el espacio columna de A, como se muestra en la figura 7.13(b). Aquí,

$$\mathbf{p} = \text{proy}_{\text{col}(A)}(\mathbf{b}) = A\overline{\mathbf{x}} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} \frac{2}{3} \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{8}{3} \\ \frac{11}{3} \end{bmatrix}$$

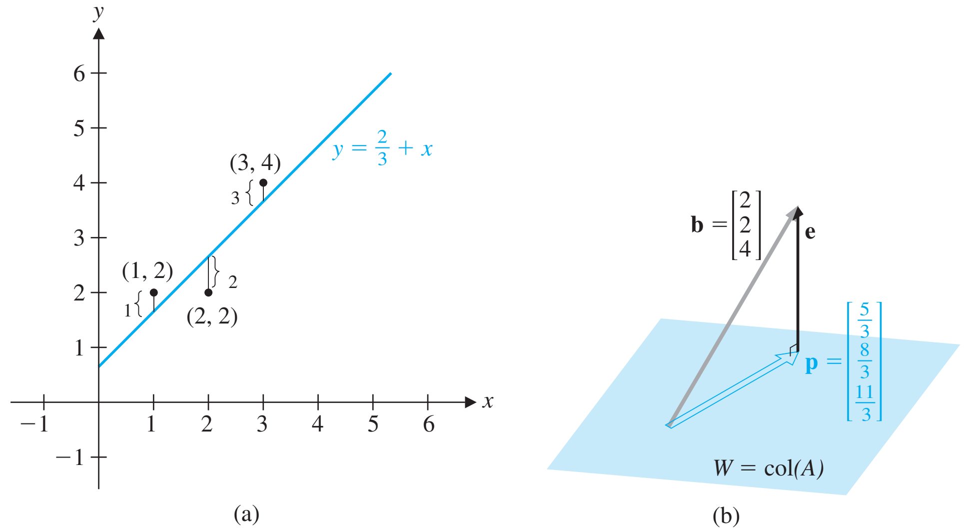


Figura 7.13

y el vector de error de mínimos cuadrados es $\mathbf{e} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}$. [¿Cómo se vería la figura 7.13(b) si los puntos de datos *fuesen* colineales?]

Ejemplo 7.27

Encuentre la aproximación lineal por mínimos cuadrados y el error de mínimos cuadrados para los puntos (1, 1), (2, 2), (3, 2) y (4, 3).

Solución Sea $y = a + bx$ la ecuación de la recta que se busca. Entonces, al sustituir los cuatro puntos en esta ecuación, se obtiene

$$\begin{aligned} a + b &= 1 \\ a + 2b &= 2 \\ a + 3b &= 2 \\ a + 4b &= 3 \end{aligned} \qquad \text{o} \qquad \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix}$$

De modo que se quiere la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix}$$

Dado que las columnas de A son linealmente independientes, la solución que se quiere es

$$\overline{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} = \begin{bmatrix} \frac{1}{2} \\ \frac{3}{5} \end{bmatrix}$$

(Compruebe este cálculo.) Por tanto, tome $a = \frac{1}{2}$ y $b = \frac{3}{5}$, lo que produce la aproximación lineal por mínimos cuadrados $y = \frac{1}{2} + \frac{3}{5}x$, como se muestra en la figura 7.14.

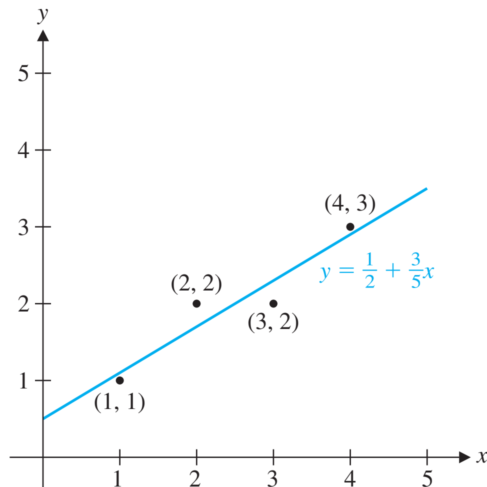
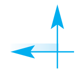


Figura 7.14

Puesto que

$$\mathbf{e} = \mathbf{b} - A\overline{\mathbf{x}} = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ \frac{3}{5} \end{bmatrix} = \begin{bmatrix} -\frac{1}{10} \\ \frac{3}{10} \\ -\frac{3}{10} \\ \frac{1}{10} \end{bmatrix}$$

el error de mínimos cuadrados es $\|\mathbf{e}\| = \sqrt{5}/5 \approx 0.447$.



Puede usar el método de mínimos cuadrados para aproximar los puntos de datos mediante curvas distintas a líneas rectas.

Ejemplo 7.28

Encuentre la parábola que ofrezca la mejor aproximación por mínimos cuadrados a los puntos $(-1, 1)$, $(0, -1)$, $(1, 0)$ y $(2, 2)$.

Solución La ecuación de una parábola es la cuadrática $y = a + bx + cx^2$. Al sustituir los puntos dados en esta cuadrática, se obtiene el sistema lineal

$$\begin{aligned} a - b + c &= 1 \\ a &= -1 \\ a + b + c &= 0 \\ a + 2b + 4c &= 2 \end{aligned} \quad \text{o} \quad \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$

Por tanto, se quiere la aproximación por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$

Calcule

$$A^TA = \begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \quad \text{y} \quad A^T\mathbf{b} = \begin{bmatrix} 2 \\ 3 \\ 9 \end{bmatrix}$$

de modo que las ecuaciones normales están dadas por

$$\begin{bmatrix} 4 & 2 & 6 \\ 2 & 6 & 8 \\ 6 & 8 & 18 \end{bmatrix} \overline{\mathbf{x}} = \begin{bmatrix} 2 \\ 3 \\ 9 \end{bmatrix}$$

cuya solución es

$$\overline{\mathbf{x}} = \begin{bmatrix} -\frac{7}{10} \\ -\frac{3}{5} \\ 1 \end{bmatrix}$$

Por tanto, la aproximación parabólica por mínimos cuadrados tiene la ecuación

$$y = -\tfrac{7}{10} - \tfrac{3}{5}x + x^2$$

como se muestra en la figura 7.15.

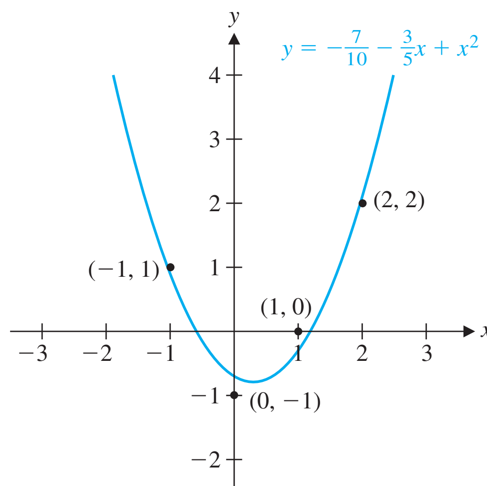


Figura 7.15
Una aproximación parabólica por mínimos cuadrados

Uno de los usos más importantes de la aproximación por mínimos cuadrados es estimar las constantes asociadas con varios procesos. El siguiente ejemplo ilustra esta aplicación en el contexto del crecimiento poblacional. Recuerde de la sección 6.7 que una población que crece (o disminuye) exponencialmente satisface una ecuación de la forma $p(t) = ce^{kt}$, donde $p(t)$ es el tamaño de la población en el tiempo t y c y k son constantes. Claramente, $c = p(0)$, pero k no es tan fácil de determinar. Es fácil ver que

$$k = \frac{p'(t)}{p(t)}$$

lo que explica por qué a k en ocasiones se le conoce como la *tasa de crecimiento relativo* de la población: es la razón entre la tasa de crecimiento $p'(t)$ y el tamaño de la población $p(t)$.

CAS **Ejemplo 7.29**

La tabla 7.2 presenta la población del mundo en intervalos de 10 años durante la segunda mitad del siglo XX. Si supone un modelo de crecimiento exponencial, encuentre la tasa de crecimiento relativa y prediga la población del mundo en 2010.

Tabla 7.2

Año	Población (en miles de millones)
1950	2.56
1960	3.04
1970	3.71
1980	4.46
1990	5.28
2000	6.08

Fuente: U.S. Bureau of the Census, International Data Base

Solución Se acuerda medir el tiempo t en intervalos de 10 años, de modo que $t = 0$ es 1950, $t = 1$ es 1960, etcétera. Dado que $c = p(0) = 2.56$, la ecuación para la tasa de crecimiento de la población es

$$p = 2.56e^{kt}$$

¿Cómo puede usar el método de mínimos cuadrados en esta ecuación? Si aplica el logaritmo natural en ambos lados, la ecuación se convierte en lineal:

$$\begin{aligned}\ln p &= \ln(2.56e^{kt})\\ &= \ln 2.56 + \ln(e^{kt})\\ &\approx 0.94 + kt\end{aligned}$$

Al colocar los valores de t y p de la tabla 7.2 se produce el siguiente sistema (donde los cálculos se redondearon a tres lugares decimales):

$$\begin{aligned}0.94 &= 0.94\\ k &= 0.172\\ 2k &= 0.371\\ 3k &= 0.555\\ 4k &= 0.724\\ 5k &= 0.865\end{aligned}$$

Puede ignorar la primera ecuación (corresponde a la condición inicial $c = p(0) = 2.56$). Las ecuaciones restantes corresponden a un sistema $A\mathbf{x} = \mathbf{b}$, con

$$A = \begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix}0.172\\0.371\\0.555\\0.724\\0.865\end{bmatrix}$$

Puesto que $A^TA = 55$ y $A^T\mathbf{b} = 9.80$, las ecuaciones normales correspondientes son justo la ecuación individual

$$55\overline{\mathbf{x}} = 9.80$$

Por tanto, $k = \overline{\mathbf{x}} = 9.80/55 \approx 0.178$. En consecuencia, la solución de mínimos cuadrados tiene la forma $p = 2.56e^{0.178t}$ (vea la figura 7.16).

La población del mundo en 2010 corresponde a $t = 6$, de donde se obtiene

$$p(6) = 2.56e^{0.178(6)} \approx 7.448$$

Por tanto, si el modelo es preciso, habrá aproximadamente 7.45 mil millones de personas en la Tierra en el año 2010. (La U.S. Census Bureau estima que la población global será "solamente" de 6.82 mil millones en 2010. ¿Por qué cree que su estimación es más alta?)

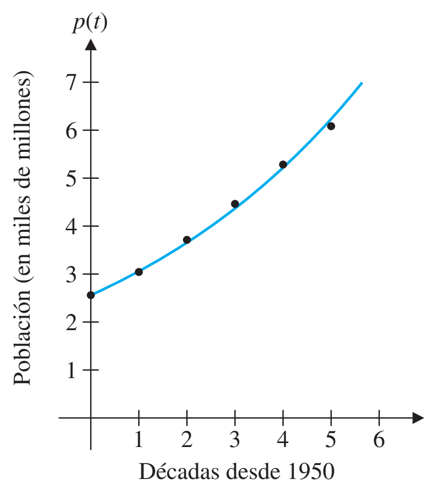


Figura 7.16

Mínimos cuadrados vía la factorización *QR*

Con frecuencia sucede que las ecuaciones normales para un problema de mínimos cuadrados están mal condicionadas. Por tanto, un pequeño error numérico al realizar eliminación gaussiana resultará en un gran error en la solución de mínimos cuadrados. En consecuencia, en la práctica, por lo general se usan otros métodos para calcular las aproximaciones por mínimos cuadrados.

Es evidente que la factorización QR de A produce una forma más confiable de calcular la aproximación por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

Teorema 7.10 Sea A una matriz de $m \times n$ con columnas linealmente independientes y sea $\mathbf{b}$ que está en $\mathbb{R}^m$. Si $A = QR$ es una factorización QR de A, entonces la única solución por mínimos cuadrados $\overline{\mathbf{x}}$ de $A\mathbf{x} = \mathbf{b}$ es

$$\overline{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$$

Demostración Recuerde del Teorema 5.16 que la factorización QR, $A = QR$ involucra una matriz Q de $m \times n$ con columnas ortonormales y una matriz triangular superior invertible R. A partir del teorema de mínimos cuadrados, se tiene

$$\begin{aligned} A^TA\overline{\mathbf{x}} &= A^T\mathbf{b} \\ \Rightarrow (QR)^TQR\overline{\mathbf{x}} &= (QR)^T\mathbf{b} \\ \Rightarrow R^TQ^TQR\overline{\mathbf{x}} &= R^TQ^T\mathbf{b} \\ \Rightarrow R^TR\overline{\mathbf{x}} &= R^TQ^T\mathbf{b} \end{aligned}$$

pues $Q^TQ = I$. (¿Por qué?)

Puesto que R es invertible, también lo es R^T, y en consecuencia se tiene

$$R\overline{\mathbf{x}} = Q^T\mathbf{b} \quad \text{o, de manera equivalente } \overline{\mathbf{x}} = R^{-1}Q^T\mathbf{b}$$

Comentario Dado que R es triangular superior, en la práctica es más sencillo resolver $R\overline{\mathbf{x}} = Q^T\mathbf{b}$ directamente que invertir R y calcular $R^{-1}Q^T\mathbf{b}$.

Ejemplo 7.30

Use la factorización QR para encontrar una solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 2 \\ -3 \\ -2 \\ 0 \end{bmatrix}$$

Solución A partir del ejemplo 5.15,

$$A = QR = \begin{bmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix}$$

Se tiene

$$Q^T\mathbf{b} = \begin{bmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ 3\sqrt{5}/10 & 3\sqrt{5}/10 & \sqrt{5}/10 & \sqrt{5}/10 \\ -\sqrt{6}/6 & 0 & \sqrt{6}/6 & \sqrt{6}/3 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ -2 \\ 0 \end{bmatrix} = \begin{bmatrix} 7/2 \\ -\sqrt{5}/2 \\ -2\sqrt{6}/3 \end{bmatrix}$$

de modo que se requiere la solución de $R\overline{\mathbf{x}} = Q^T\mathbf{b}$, o

$$\begin{bmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{bmatrix} \overline{\mathbf{x}} = \begin{bmatrix} 7/2 \\ -\sqrt{5}/2 \\ -2\sqrt{6}/3 \end{bmatrix}$$

La sustitución inversa produce rápidamente

$$\overline{\mathbf{x}} = \begin{bmatrix} 4/3 \\ 3/2 \\ -4/3 \end{bmatrix}$$

Revisión de la proyección ortogonal

Uno de los subproductos amigables del método de mínimos cuadrados es una nueva fórmula para la proyección ortogonal de un vector sobre un subespacio de $\mathbb{R}^m$.

Teorema 7.11

Sea W un subespacio de $\mathbb{R}^m$ y sea A una matriz de $m \times n$ cuyas columnas forman una base para W. Si $\mathbf{v}$ es cualquier vector en $\mathbb{R}^m$, entonces la proyección ortogonal de $\mathbf{v}$ sobre W es el vector

$$\text{proy}_W(\mathbf{v}) = A(A^TA)^{-1}A^T\mathbf{v}$$

La transformación lineal $P : \mathbb{R}^m \to \mathbb{R}^m$ que proyecta $\mathbb{R}^m$ sobre W tiene $A(A^TA)^{-1}A^T$ como su matriz estándar.

Demostración Dada la forma como se construyó A, su espacio columna es W. Puesto que las columnas de A son linealmente independientes, el teorema de mínimos cuadrados garantiza que existe una solución única de mínimos cuadrados para $A\mathbf{x} = \mathbf{v}$ dada por

$$\overline{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{v}$$

Por la ecuación (1),

$$A\bar{\mathbf{x}} = \text{proy}_{\text{col}(A)}(\mathbf{v}) = \text{proy}_W(\mathbf{v})$$

Por tanto, $\text{proy}_W(\mathbf{v}) = A((A^TA)^{-1}A^T\mathbf{v}) = (A(A^TA)^{-1}A^T)\mathbf{v}$

como se requería. Dado que esta ecuación se cumple para todo $\mathbf{v}$ en $\mathbb{R}^m$, el último enunciado del teorema se obtiene inmediatamente.

El Teorema 7.11 se ilustrará al revisar nuevamente el ejemplo 5.11.

Ejemplo 7.31

Encuentre la proyección ortogonal de $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix}$ sobre el plano W en $\mathbb{R}^3$ con ecuación $x - y + 2z = 0$, y proporcione la matriz estándar de la transformación de proyección ortogonal sobre W.

Solución Como en el ejemplo 5.11 se tomará como base para W el conjunto

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Se forma la matriz

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$$

con estos vectores base como sus columnas. Entonces

$$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

de modo que $$(A^TA)^{-1} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix}$$

Por el Teorema 7.11, la matriz estándar de la transformación de proyección ortogonal sobre W es

$$A(A^TA)^{-1}A^T = A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} & -\frac{1}{3} \\ \frac{1}{6} & \frac{5}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

de modo que la proyección ortogonal de $\mathbf{v}$ sobre W es

$$\text{proy}_W(\mathbf{v}) = A(A^TA)^{-1}A^T\mathbf{v} = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} & -\frac{1}{3} \\ \frac{1}{6} & \frac{5}{6} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix} \begin{bmatrix} 3 \\ -1 \\ 2 \end{bmatrix} = \begin{bmatrix} \frac{5}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \end{bmatrix}$$

que concuerda con la solución al ejemplo 5.11.



Comentario Dado que la proyección de un vector sobre un subespacio W es única, la matriz estándar de esta transformación lineal (dada por el Teorema 7.11) no puede depender de la elección de base para W. En otras palabras, con una base diferente para W, se tiene una matriz A diferente, ¡pero la matriz $A(A^TA)^{-1}A^T$ será la misma! (En el ejercicio 43 se le pide verificar esto.)

La pseudoinversa de una matriz

Si A es una matriz de $n \times n$ con columnas linealmente independientes, entonces es invertible, y la única solución de $A\mathbf{x} = \mathbf{b}$ es $\mathbf{x} = A^{-1}\mathbf{b}$. Si $m > n$ y A es de $m \times n$ con columnas linealmente independientes, entonces $A\mathbf{x} = \mathbf{b}$ no tiene solución exacta, pero la mejor aproximación está dada por la solución única de mínimos cuadrados $\overline{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$. En consecuencia, la matriz $(A^TA)^{-1}A^T$ desempeña el papel de un "inverso de A" en esta situación.

Definición Si A es una matriz con columnas linealmente independientes, entonces la ***pseudoinversa*** de A es la matriz A^+ definida por

$$A^+ = (A^TA)^{-1}A^T$$

Observe que si A es de $m \times n$, entonces A^+ es de $n \times m$.

Ejemplo 7.32

Encuentre la pseudoinversa de $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}$.

Solución En el ejemplo 7.26 ya se hicieron la mayoría de los cálculos. Al usar el trabajo previo se tiene

$$A^+ = (A^TA)^{-1}A^T = \begin{bmatrix} \frac{7}{3} & -1 \\ -1 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3} & \frac{1}{3} & -\frac{2}{3} \\ -\frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$$

La pseudoinversa es una notación abreviada conveniente para algunos de los conceptos que se han explorado. Por ejemplo, si A es de $m \times n$, con columnas linealmente independientes, la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ está dada por

$$\overline{\mathbf{x}} = A^+\mathbf{b}$$

y la matriz estándar de la proyección ortogonal P de $\mathbb{R}^m$ sobre col(A) es

$$[P] = AA^+$$

Si A en realidad es una matriz cuadrada, entonces es fácil demostrar que $A^+ = A^{-1}$ (vea el ejercicio 53). En este caso, la solución de mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ es la solución *exacta*, pues

$$\overline{\mathbf{x}} = A^+\mathbf{b} = A^{-1}\mathbf{b} = \mathbf{x}$$


La matriz de proyección se convierte en $[P] = AA^+ = AA^{-1} = I$. (¿Cuál es la interpretación geométrica de esta igualdad?)


El teorema 7.12 resume las propiedades clave de la pseudoinversa de una matriz. (Antes de leer la demostración de este teorema, verifique estas propiedades para la matriz en el ejemplo 7.32.)

Teorema 7.12 Sea A una matriz con columnas linealmente independientes. Entonces la pseudoinversa A^+ de A satisface las siguientes propiedades, llamadas ***condiciones de Penrose*** para A:

a. $AA^+A = A$
b. $A^+AA^+ = A^+$
c. AA^+ y A^+A son simétricas.

Demostración Se demuestra la condición (a) y la mitad de la condición (c), y se dejan las demostraciones de las condiciones restantes como los ejercicios 54 y 55.

(a) Calcule

$$\begin{aligned} AA^+A &= A((A^TA)^{-1}A^T)A \\ &= A(A^TA)^{-1}(A^TA) \\ &= AI = A \end{aligned}$$

(c) Por el Teorema 3.4, A^TA es simétrica. Por tanto, $(A^TA)^{-1}$ también es simétrica, por el ejercicio 46 de la sección 3.3. Al tomar la transpuesta de AA^+ se tiene

$$\begin{aligned} (AA^+)^T &= (A(A^TA)^{-1}A^T)^T \\ &= (A^T)^T((A^TA)^{-1})^TA^T \\ &= A(A^TA)^{-1}A^T \\ &= AA^+ \end{aligned}$$

El ejercicio 56 explora más propiedades de la pseudoinversa. En la siguiente sección se verá cómo extender la definición de A^+ para manejar *todas* las matrices, ya sea que las columnas de A sean o no linealmente independientes.

Ejercicios 7.3

CAS

En los ejercicios 1-3, considere los puntos de datos (1, 0), (2, 1) *y* (3, 5). *Calcule el error de mínimos cuadrados para la recta dada. En cada caso, grafique los puntos y la recta.*

1. $y = -2 + 2x$ **2.** $y = -3 + 2x$ **3.** $y = -3 + \frac{5}{2}x$

En los ejercicios 4-6, considere los puntos de datos $(-5, 3)$, (0, 3), (5, 2) *y* (10, 0). *Calcule el error de mínimos cuadrados para la recta dada. En cada caso, grafique los puntos y la recta.*

4. $y = 2 - x$ **5.** $y = \frac{5}{2}$ **6.** $y = 2 - \frac{1}{5}x$

En los ejercicios 7-14, encuentre la aproximación lineal por mínimos cuadrados para los puntos dados y calcular el correspondiente error de mínimos cuadrados.

7. (1, 0), (2, 1), (3, 5)
8. (1, 5), (2, 3), (3, 2)
9. (0, 4), (1, 1), (2, 0)
10. (0, 2), (1, 2), (2, 5)
11. $(-5, -1)$, (0, 1), (5, 2), (10, 4)

12. $(-5, 3), (0, 3), (5, 2), (10, 0)$

13. $(1, 1), (2, 3), (3, 4), (4, 5), (5, 7)$

14. $(1, 10), (2, 8), (3, 5), (4, 3), (5, 0)$

En los ejercicios 15-18, encuentre la aproximación parabólica por mínimos cuadrados para los puntos dados.

15. $(1, 1), (2, -2), (3, 3), (4, 4)$

16. $(1, 8), (2, 7), (3, 5), (4, 2)$

17. $(-2, 4), (-1, 7), (0, 3), (1, 0), (2, -1)$

18. $(-2, 0), (-1, -11), (0, -10), (1, -9), (2, 8)$

En los ejercicios 19-22, encuentre una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ al construir y resolver las ecuaciones normales.

19. $A = \begin{bmatrix} 3 & 1 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

20. $A = \begin{bmatrix} 3 & -2 \\ 1 & -2 \\ 2 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

21. $A = \begin{bmatrix} 1 & -2 \\ 0 & -3 \\ 2 & 5 \\ 3 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 4 \\ 1 \\ -2 \\ 4 \end{bmatrix}$

22. $A = \begin{bmatrix} 2 & 0 \\ 1 & -1 \\ 3 & 1 \\ -1 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 1 \\ -1 \\ 3 \end{bmatrix}$

En los ejercicios 23 y 24, demuestre que la solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$ no es única y resuelva las ecuaciones normales para encontrar todas las soluciones de mínimos cuadrados.

23. $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & 1 \\ 1 & -1 & 1 & 0 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ -3 \\ 2 \\ 4 \end{bmatrix}$

24. $A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & -1 & 1 & -1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 5 \\ 3 \\ -1 \\ 1 \end{bmatrix}$

En los ejercicios 25 y 26, encuentre la mejor aproximación a una solución del sistema de ecuaciones dado.

25. $\begin{aligned} x + y - z &= 2 \\ -y + 2z &= 6 \\ 3x + 2y - z &= 11 \\ -x + \qquad z &= 0 \end{aligned}$

26. $\begin{aligned} 2x + 3y + z &= 21 \\ x + y + z &= 7 \\ -x + y - z &= 14 \\ 2y + z &= 0 \end{aligned}$

En los ejercicios 27 y 28, se proporciona una factorización QR de A. Úsela para encontrar una solución por mínimos cuadrados de $A\mathbf{x} = \mathbf{b}$.

27. $A = \begin{bmatrix} 2 & 1 \\ 2 & 0 \\ 1 & 1 \end{bmatrix}, Q = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}, R = \begin{bmatrix} 3 & 1 \\ 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$

28. $A = \begin{bmatrix} 1 & 0 \\ 2 & -1 \\ -1 & 1 \end{bmatrix}, Q = \begin{bmatrix} 1/\sqrt{6} & 1/\sqrt{2} \\ 2/\sqrt{6} & 0 \\ -1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix},$

$R = \begin{bmatrix} \sqrt{6} & -\sqrt{6}/2 \\ 0 & 1/\sqrt{2} \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

29. Una pelota de tenis se suelta desde varias alturas y se mide la altura que alcanza la pelota en el primer rebote. Use los datos de la tabla 7.3 para encontrar la aproximación lineal por mínimos cuadrados para la altura de rebote b como una función lineal de la altura inicial h.

Tabla 7.3

***h* (cm)**	20	40	48	60	80	100
***b* (cm)**	14.5	31	36	45.5	59	73.5

30. La ley de Hooke afirma que la longitud L de un resorte es una función lineal de la fuerza F que se le aplica. (Vea la figura 7.17 y el ejemplo 6.92.) En concordancia, existen constantes a y b tales que

$$L = a + bF$$

La tabla 7.4 muestra los resultados de colocar varios pesos en un resorte.

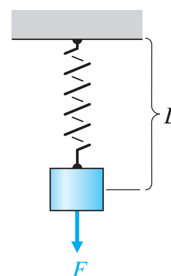


Figura 7.17

Tabla 7.4

***F* (oz)**	2	4	6	8
***L* (in.)**	7.4	9.6	11.5	13.6

Tabla 7.5

Año de nacimiento	1920	1930	1940	1950	1960	1970	1980	1990
Esperanza de vida (años)	54.1	59.7	62.9	68.2	69.7	70.8	73.7	75.4

Fuente: World Almanac and Book of Facts. Nueva York: World Almanac Books, 1999

(a) Determine las constantes a y b al encontrar la aproximación lineal por mínimos cuadrados para dichos datos. ¿Qué representa a?
(b) Estime la longitud del resorte cuando se le coloca un peso de 5 onzas.

31. La tabla 7.5 proporciona las esperanzas de vida para las personas nacidas en Estados Unidos en los años dados.
(a) Determine la aproximación lineal por mínimos cuadrados para dichos datos y úselos para predecir la esperanza de vida de alguien nacido en 2000.
(b) ¿Qué tan bueno es este modelo? Explique.

32. Cuando un objeto se lanza recto hacia arriba en el aire, la segunda ley de movimiento de Newton afirma que su altura $s(t)$ en el tiempo t está dado por

$$s(t) = s_0 + v_0 t + \tfrac{1}{2}gt^2$$

donde v_0 es su velocidad inicial y g es la constante de aceleración debida a la gravedad. Suponga que se toman las mediciones que se muestran en la tabla 7.6.

Tabla 7.6

Tiempo (s)	0.5	1	1.5	2	3
Altura (m)	11	17	21	23	18

(a) Encuentre la aproximación cuadrática por mínimos cuadrados para dichos datos.
(b) Estime la altura a la cual se liberó el objeto (en m), su velocidad inicial (en m/s) y su aceleración debida a la gravedad (en m/s^2).
(c) ¿Aproximadamente cuándo golpeará el suelo el objeto?

33. La tabla 7.7 proporciona la población de Estados Unidos a intervalos de 10 años para los años 1950-2000.
(a) Si supone un modelo de crecimiento exponencial de la forma $p(t) = ce^{kt}$, donde $p(t)$ es la población en el tiempo t, use mínimos cuadrados para encontrar la ecuación para la tasa de crecimiento de la población. [*Sugerencia:* sea $t = 0$ en 1950.]
(b) Use la ecuación para estimar la población estadounidense en 2010.

Tabla 7.7

Año	Población (en millones)
1950	150
1960	179
1970	203
1980	227
1990	250
2000	281

Fuente: U.S. Bureau of the Census

34. La tabla 7.8 muestra los salarios promedio de las grandes ligas de beisbol para los años 1970-2005.
(a) Encuentre la aproximación cuadrática por mínimos cuadrados para dichos datos.
(b) Encuentre la aproximación exponencial por mínimos cuadrados para dichos datos.
(c) ¿Cuál ecuación brinda la mejor aproximación? ¿Por qué?
(d) ¿Cuál estima que será el salario promedio en las grandes ligas de beisbol en 2010 y 2015?

Tabla 7.8

Año	Salario promedio (miles de dólares)
1970	29.3
1975	44.7
1980	143.8
1985	371.6
1990	597.5
1995	1110.8
2000	1895.6
2005	2476.6

Fuente: Major League Baseball Players Association

35. Una muestra de 200 mg de polonio 210 radiactivo se observa mientras decae. La tabla 7.9 muestra la masa restante en varios momentos.

Si supone un modelo de decaimiento exponencial, use mínimos cuadrados para encontrar la vida media del polonio 210. (Vea la sección 6.7.)

Tabla 7.9

Tiempo (días)	0	30	60	90
Masa (mg)	200	172	148	128

36. Encuentre el plano $z = a + bx + cy$ que ajusta mejor los puntos de datos $(0, -4, 0)$, $(5, 0, 0)$, $(4, -1, 1)$, $(1, -3, 1)$, y $(-1, -5, -2)$.

En los ejercicios 37-42, encuentre la matriz estándar de la proyección ortogonal sobre el subespacio W. Luego use esta matriz para encontrar la proyección ortogonal de **v** *sobre W.*

37. $W = \text{gen}\left(\begin{bmatrix}1\\1\end{bmatrix}\right), \mathbf{v} = \begin{bmatrix}3\\4\end{bmatrix}$

38. $W = \text{gen}\left(\begin{bmatrix}1\\-2\end{bmatrix}\right), \mathbf{v} = \begin{bmatrix}1\\1\end{bmatrix}$

39. $W = \text{gen}\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right), \mathbf{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}$

40. $W = \text{gen}\left(\begin{bmatrix}2\\2\\-1\end{bmatrix}\right), \mathbf{v} = \begin{bmatrix}1\\0\\0\end{bmatrix}$

41. $W = \text{gen}\left(\begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right), \mathbf{v} = \begin{bmatrix}1\\0\\0\end{bmatrix}$

42. $W = \text{gen}\left(\begin{bmatrix}1\\-2\\1\end{bmatrix}, \begin{bmatrix}1\\0\\-1\end{bmatrix}\right), \mathbf{v} = \begin{bmatrix}1\\2\\3\end{bmatrix}$

43. Verifique que la matriz estándar de la proyección sobre W en el ejemplo 7.31 (como se construye mediante el Teorema 7.11) no depende de la elección de base. Tome

$$\left\{\begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\3\\1\end{bmatrix}\right\}$$

como una base para W y repita los cálculos para demostrar que la matriz de proyección resultante es la misma.

44. Sea A una matriz con columnas linealmente independientes y sea $P = A(A^TA)^{-1}A^T$ la matriz de proyección ortogonal sobre col(A).

(a) Demuestre que P es simétrica.

(b) Demuestre que P es idempotente.

En los ejercicios 45-52, calcule la pseudoinversa de A.

45. $A = \begin{bmatrix}1\\2\end{bmatrix}$

46. $A = \begin{bmatrix}1\\-1\\2\end{bmatrix}$

47. $A = \begin{bmatrix}1 & 3\\-1 & 1\\0 & 2\end{bmatrix}$

48. $A = \begin{bmatrix}1 & 3\\3 & 1\\2 & 2\end{bmatrix}$

49. $A = \begin{bmatrix}1 & 1\\0 & 1\end{bmatrix}$

50. $A = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$

51. $A = \begin{bmatrix}1 & 0 & 0\\1 & 0 & 1\\0 & 1 & 1\\1 & 1 & 1\end{bmatrix}$

52. $A = \begin{bmatrix}1 & 2 & 0\\0 & 1 & -1\\1 & 1 & -2\\0 & 0 & 2\end{bmatrix}$

53. (a) Demuestre que, si A es una matriz cuadrada con columnas linealmente independientes, entonces $A^+ = A^{-1}$.

(b) Si A es una matriz de $m \times n$ con columnas ortonormales, ¿cuál es A^+?

54. Demuestre el Teorema 7.12(b).

55. Demuestre la parte restante del Teorema 7.12(c).

56. Sea A una matriz con columnas linealmente independientes. Demuestre lo siguiente:

(a) $(cA)^+ = (1/c)A^+$ para todos los escalares $c \neq 0$.

(b) $(A^+)^+ = A$ si A es una matriz cuadrada.

(c) $(A^T)^+ = (A^+)^T$ si A es una matriz cuadrada.

57. Se proporcionan n puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$. Demuestre que si los puntos no yacen todos en la misma recta vertical, entonces tienen una aproximación lineal única por mínimos cuadrados.

58. Se proporcionan n puntos de datos $(x_1, y_1), \ldots, (x_n, y_n)$. Generalice el ejercicio 57 para demostrar que si al menos $k + 1$ de los $x_1, \ldots, x_n$ son distintos, entonces los puntos dados tienen una aproximación polinomial única por mínimos cuadrados de grado máximo k.

7.4 La descomposición de valor singular

En el capítulo 5 se vio que toda matriz simétrica A puede factorizarse como $A = PDP^T$, donde P es una matriz ortogonal y D es una matriz diagonal que muestra los eigenvalores de A. Si A no es simétrica, no es posible tal factorización, pero, como se aprendió en el capítulo 4, todavía es posible factorizar una matriz cuadrada A como $A = PDP^{-1}$, donde D es como antes, pero P ahora es simplemente una matriz invertible. Sin embargo, no toda matriz es diagonalizable, de modo que puede sorprenderle que ahora se demostrará que *toda* matriz (simétrica o no, cuadrada o no) tiene una factorización de la forma $A = PDQ^T$, ¡donde P y Q son ortogonales y D es una matriz diagonal! Este resultado notable es la *descomposición de valor singular* (DVS) y es una de las factorizaciones más importantes de todas las matrices.

En esta sección se demostrará cómo calcular la DVS de una matriz y luego considerar algunas de sus muchas aplicaciones. En el trayecto se atarán algunos cabos sueltos y se responderán algunas preguntas que quedaron planteadas en secciones anteriores.

Los valores singulares de una matriz

Para cualquier matriz A de $m \times n$, la matriz A^TA de $n \times n$ es simétrica y en consecuencia puede ser diagonalizable ortogonalmente , por el teorema espectral. No sólo todos los eigenvalores de A^TA son reales (Teorema 5.18), todos son *no negativos*. Para demostrar esto, sea λ un eigenvalor de A^TA con su correspondiente eigenvector unitario $\mathbf{v}$. Entonces

$$\begin{aligned} 0 \le \|A\mathbf{v}\|^2 &= (A\mathbf{v})\cdot(A\mathbf{v}) = (A\mathbf{v})^TA\mathbf{v} = \mathbf{v}^TA^TA\mathbf{v} \\ &= \mathbf{v}^T\lambda\mathbf{v} = \lambda(\mathbf{v}\cdot\mathbf{v}) = \lambda\|\mathbf{v}\|^2 = \lambda \end{aligned}$$

Por tanto, tiene sentido sacar raíces cuadradas (positivas) de dichos eigenvalores.

Definición Si A es una matriz de $m \times n$, los ***valores singulares*** de A son las raíces cuadradas de los eigenvalores de A^TA y se denotan mediante $\sigma_1, \ldots, \sigma_n$. Es convencional ordenar los valores singulares de modo que $\sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_n$.

Ejemplo 7.33 Encuentre los valores singulares de

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Solución La matriz

$$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

tiene eigenvalores $\lambda_1 = 3$ y $\lambda_2 = 1$. En consecuencia, los valores singulares de A son $\sqrt{\lambda_1} = \sqrt{3}$ y $\sigma_2 = \sqrt{\lambda_2} = 1$.

Para comprender el significado de los valores singulares de una matriz A de $m \times n$, considere los eigenvectores de A^TA. Puesto que A^TA es simétrica, se sabe que existe una base *ortonormal* para $\mathbb{R}^n$ que consiste de los eigenvectores de A^TA. Sea $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ una base tal que corresponde a los eigenvalores de A^TA, ordenados de modo que $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. A partir de los cálculos justo antes de la definición,

$$\lambda_i = \|A\mathbf{v}_i\|^2$$

Por tanto,

$$\sigma_i = \sqrt{\lambda_i} = \|A\mathbf{v}_i\|$$

En otras palabras, los valores singulares de A son las longitudes de los vectores $A\mathbf{v}_1, \ldots, A\mathbf{v}_n$.

Geométricamente, este resultado tiene una amigable interpretación. Considere nuevamente el ejemplo 7.33. Si $\mathbf{x}$ se encuentra sobre la circunferencia unitaria en $\mathbb{R}^2$ (esto es $\|\mathbf{x}\| = 1$), entonces

$$\begin{aligned}\|A\mathbf{x}\|^2 &= (A\mathbf{x}) \cdot (A\mathbf{x}) = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^TA^TA\mathbf{x} \\ &= \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + 2x_1x_2 + 2x_2^2\end{aligned}$$

que se reconoce como una forma cuadrática. Por el Teorema 5.25, los valores máximo y mínimo de esta forma cuadrática, sujetos a la restricción, $\|\mathbf{x}\| = 1$, son $\lambda_1 = 3$ y $\lambda_2 = 1$, respectivamente, y se presentan en los correspondientes eigenvectores de A^TA; esto es, cuando $\mathbf{x} = \mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ y $\mathbf{x} = \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$, respectivamente. Puesto que

$$\|A\mathbf{v}_i\|^2 = \mathbf{v}_i^TA^TA\mathbf{v}_i = \lambda_i$$

para $i = 1, 2$, se ve que $\sigma_1 = \|A\mathbf{v}_1\| = \sqrt{3}$ y $\sigma_2 = \|A\mathbf{v}_2\| = 1$ son los valores máximo y mínimo de las longitudes $\|A\mathbf{x}\|$ conforme $\mathbf{x}$ recorre la circunferencia unitaria en $\mathbb{R}^2$.

Ahora, la transformación lineal que corresponde a A mapea $\mathbb{R}^2$ sobre el plano en $\mathbb{R}^3$



con ecuación $x - y - z = 0$ (verifique esto), y la imagen de la circunferencia unitaria bajo esta transformación es una elipse que se encuentra en este plano. (Dentro de poco se verificará este hecho en general; vea la figura 7.18.) De modo que σ_1 y σ_2 son las longitudes de la mitad de los ejes mayor y menor de esta elipse, como se muestra en la figura 7.19.

Ahora es posible describir la descomposición de valor singular de una matriz.

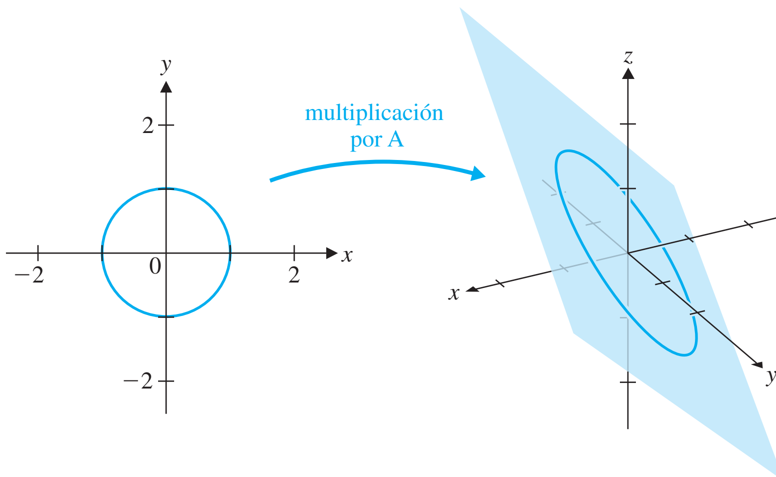


Figura 7.18
La matriz A transforma la circunferencia unitaria en $\mathbb{R}^2$ en una elipse en $\mathbb{R}^3$.

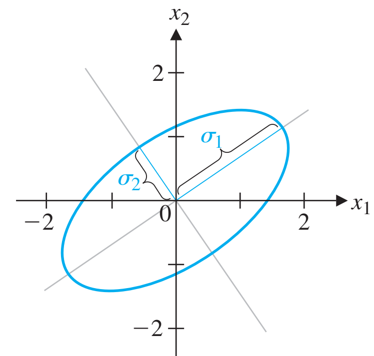


Figura 7.19

La descomposición de valor singular

Se quiere demostrar que una matriz A de $m \times n$ puede factorizarse como

$$A = U\Sigma V^T$$

donde U es una matriz ortogonal de $m \times m$, V es una matriz ortogonal de $n \times n$ y Σ es una matriz "diagonal" de $m \times n$. Si los valores singulares *distintos de cero* de A son

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

y $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$, entonces Σ tendrá la forma de bloque

$$\Sigma = \begin{bmatrix} \overbrace{D}^{r} & \overbrace{O}^{n-r} \\ O & O \end{bmatrix} \begin{matrix} \}\, r \\ \}\, m-r \end{matrix}, \quad \text{donde} \quad D = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix} \tag{1}$$

y cada matriz O es una matriz cero del tamaño adecuado. (Si $r = m$ o $r = n$, alguna de éstas no aparecerá.) Algunos ejemplos de tal matriz Σ con $r = 2$ son

$$\Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$



(¿Cuál es D en cada caso?)

Para construir la matriz ortogonal V, primero encuentre una base ortonormal $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ para $\mathbb{R}^n$ que consiste de eigenvectores de la matriz simétrica A^TA de $n \times n$. Entonces

$$V = [\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_n]$$

es una matriz ortogonal de $n \times n$.

Para la matriz ortogonal U, note primero que $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_n\}$ es un conjunto ortogonal de vectores en $\mathbb{R}^m$. Para ver esto, suponga que $\mathbf{v}_i$ es el eigenvector de A^TA correspondiente al eigenvalor λ_i. Entonces, para $i \neq j$, se tiene

$$\begin{aligned}(A\mathbf{v}_i) \cdot (A\mathbf{v}_j) &= (A\mathbf{v}_i)^T A\mathbf{v}_j \\ &= \mathbf{v}_i^T A^T A\mathbf{v}_j \\ &= \mathbf{v}_i^T \lambda_j \mathbf{v}_j \\ &= \lambda_j(\mathbf{v}_i \cdot \mathbf{v}_j) = 0\end{aligned}$$

dado que los eigenvectores $\mathbf{v}_i$ son ortogonales. Ahora recuerde que los valores singulares satisfacen $\sigma_i = \|A\mathbf{v}_i\|$ y que los primeros r de éstos son distintos de cero. Por tanto, puede normalizar $A\mathbf{v}_1, \ldots, A\mathbf{v}_r$ al establecer

$$\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i \quad \text{para } i = 1, \ldots, r$$

Esto garantiza que $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ es un conjunto ortonormal en $\mathbb{R}^m$, pero si $r < m$ no será una base para $\mathbb{R}^m$. En este caso, se extiende el conjunto $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ a una base ortonormal $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ para $\mathbb{R}^m$. (Esta es la única parte complicada de la construcción; en los ejemplos siguientes y en los ejercicios se describirán técnicas para realizarla.) Entonces se establece

$$U = [\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_m]$$

Todo lo que resta por demostrar es que esto funciona; es decir, es necesario verificar que con U, V y Σ como se describieron, se tiene $A = U\Sigma V^T$. Puesto que $V^T = V^{-1}$, esto equivale a demostrar que

$$AV = U\Sigma$$

Se sabe que $$A\mathbf{v}_i = \sigma_i \mathbf{u}_i \quad \text{para } i = 1, \ldots, r$$

y $\|A\mathbf{v}_i\| = \sigma_i = 0$ para $i = r + 1, \ldots, n$. En consecuencia,

$$A\mathbf{v}_i = \mathbf{0} \quad \text{para } i = r + 1, \ldots, n$$

Por tanto,
$$\begin{aligned} AV &= A[\mathbf{v}_1 \ \cdots \ \mathbf{v}_n] \\ &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_n] \\ &= [A\mathbf{v}_1 \ \cdots \ A\mathbf{v}_r \ \mathbf{0} \ \cdots \ \mathbf{0}] \\ &= [\sigma_1\mathbf{u}_1 \ \cdots \ \sigma_r\mathbf{u}_r \ \mathbf{0} \ \cdots \ \mathbf{0}] \\ &= [\mathbf{u}_1 \ \cdots \ \mathbf{u}_m] \left[\begin{array}{ccc|c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hline & O & & O \end{array}\right] \\ &= U\Sigma \end{aligned}$$

como se requería.

Acaba de demostrarse el siguiente teorema extremadamente importante.

Teorema 7.13 **La descomposición de valor singular**

Sea A una matriz de $m \times n$ con valores singulares $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ y $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$. Entonces existe una matriz ortogonal U de $m \times m$, una matriz ortogonal V de $n \times n$ y una matriz Σ de $m \times n$ de la forma que se muestra en la ecuación (1) tal que

$$A = U\Sigma V^T$$

Una factorización de A como en el Teorema 7.13 se llama ***descomposición de valor singular (DVS)*** de A. Las columnas de U se llaman ***vectores singulares izquierdos*** de A, y las columnas de V se llaman ***vectores singulares derechos*** de A. Las matrices U y V no están determinadas exclusivamente por A, sino que Σ debe contener los valores singulares de A, como en la ecuación (1). (Vea el ejercicio 25.)

Ejemplo 7.34 Encuentre una descomposición de valor singular para las siguientes matrices:

(a) $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

Solución (a) Calcule

$$A^TA = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

y descubra que sus eigenvalores son $\lambda_1 = 2$, $\lambda_2 = 1$ y $\lambda_3 = 0$, con sus correspondientes eigenvectores

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$



(Verifique esto.) Estos vectores son ortogonales, de modo que se les normaliza para obtener

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$$

Los valores singulares de A son $\sigma_1 = \sqrt{2}$, $\sigma_2 = \sqrt{1} = 1$, y $\sigma_3 = \sqrt{0} = 0$. Por tanto,

$$V = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} \quad \text{y} \quad \Sigma = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Para encontrar U, calcule

$$\mathbf{u}_1 = \frac{1}{\sigma_1}A\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

y

$$\mathbf{u}_2 = \frac{1}{\sigma_2}A\mathbf{v}_2 = \frac{1}{1}\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Estos dichos vectores ya forman una base ortonormal (la base estándar) para $\mathbb{R}^2$, así que se tiene

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Esto produce la DVS

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} = U\Sigma V^T$$



que puede comprobarse fácilmente. (Note que V tiene que ser transpuesta. Note también que el valor singular σ_3 no aparece en Σ.)

(b) Esta es la matriz del ejemplo 7.33, así que ya se sabe que los valores singulares son $\sigma_1 = \sqrt{3}$ y $\sigma_2 = 1$, lo que corresponde a $\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ y $\mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$. De modo que

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Para U, calcule

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

y

$$\mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \frac{1}{1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Esta vez es necesario extender $\{\mathbf{u}_1, \mathbf{u}_2\}$ a una base ortonormal para $\mathbb{R}^3$. Existen varias formas de proceder; un método es usar el proceso de Gram-Schmidt, como en el ejemplo 5.14. Primero es necesario encontrar un conjunto linealmente independiente de tres vectores que contengan a $\mathbf{u}_1$ y $\mathbf{u}_2$. Si $\mathbf{e}_3$ es el tercer vector base estándar en $\mathbb{R}^3$, es claro que $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_3\}$ es linealmente independiente. (Aquí debe determinar esto por inspección, pero un método confiable para usar en general es reducir por renglón la matriz con dichos vectores como sus columnas y usar el teorema fundamental.) Al aplicar Gram-Schmidt (con normalización) a $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_3\}$ (sólo es necesario el último paso), se encuentra

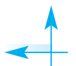
$$\mathbf{u}_3 = \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

de modo que

$$U = \begin{bmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$$

y se tiene la DVS

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = U\Sigma V^T$$



Existe otra forma de la descomposición de valor singular, análoga a la descomposición espectral de una matriz simétrica. Se obtiene a partir de la DVS mediante una expansión por producto externo y es muy útil en las aplicaciones. Puede obtener esta versión de la DVS al imitar lo que se hizo para obtener la descomposición espectral.

En concordancia, se tiene

$$
\begin{aligned}
A = U\Sigma V^T &= [\mathbf{u}_1 \;\cdots\; \mathbf{u}_m]\left[\begin{array}{ccc:c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hdashline & O & & O \end{array}\right]\begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} \\
&= [\mathbf{u}_1 \;\cdots\; \mathbf{u}_r \,\vdots\, \mathbf{u}_{r+1} \;\cdots\; \mathbf{u}_m]\left[\begin{array}{ccc:c} \sigma_1 & \cdots & 0 & \\ \vdots & \ddots & \vdots & O \\ 0 & \cdots & \sigma_r & \\ \hdashline & O & & O \end{array}\right]\left[\begin{array}{c} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_r^T \\ \hdashline \mathbf{v}_{r+1}^T \\ \vdots \\ \mathbf{v}_n^T \end{array}\right] \\
&= [\mathbf{u}_1 \;\cdots\; \mathbf{u}_r]\begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix}\begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_r^T \end{bmatrix} + [\mathbf{u}_{r+1} \;\cdots\; \mathbf{u}_m][O]\begin{bmatrix} \mathbf{v}_{r+1}^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix} \\
&= [\mathbf{u}_1 \;\cdots\; \mathbf{u}_r]\begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{bmatrix}\begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_r^T \end{bmatrix} \\
&= [\sigma_1\mathbf{u}_1 \;\cdots\; \sigma_r\mathbf{u}_r]\begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_r^T \end{bmatrix} \\
&= \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T
\end{aligned}
$$

al usar la multiplicación en bloque y la representación columna-renglón del producto. El siguiente teorema resume el proceso para obtener esta ***forma de producto externo de la DVS***.

Teorema 7.14 La forma de producto externo de la DVS

Sea A una matriz de $m \times n$ con valores singulares $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ y $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$. Sean $\mathbf{u}_1, \ldots, \mathbf{u}_r$ vectores singulares izquierdos y sean $\mathbf{v}_1, \ldots, \mathbf{v}_r$ vectores singulares derechos de A que corresponden a dichos valores singulares. Entonces

$$A = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T$$

Comentario Si A es una matriz simétrica positiva definida, entonces los Teoremas 7.13 y 7.14 se reducen a resultados que ya se conocen. En este caso, no es difícil demostrar que la DVS generaliza el teorema espectral y que el Teorema 7.14 generaliza la descomposición espectral. (Vea el ejercicio 27.)

La DVS de una matriz A contiene mucha información importante acerca de A, como se destaca en el Teorema 7.15, que es crucial.

Teorema 7.15

Sea $A = U\Sigma V^T$ una descomposición de valor singular de una matriz A de $m \times n$. Sean $\sigma_1, \ldots, \sigma_r$ todos los valores singulares distintos de cero de A. Entonces:

a. El rank de A es r.
b. $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ es una base ortonormal para col(A).
c. $\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\}$ es una base ortonormal para nulo(A^T).
d. $\{\mathbf{v}_1, \ldots, \mathbf{v}_r\}$ es una base ortonormal para renglón(A).
e. $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ es una base ortonormal para nulo(A).

Demostración (a) Por el ejercicio 61 de la sección 3.5, se tiene

$$\begin{aligned} \text{rank}(A) &= \text{rank}(U\Sigma V^T) \\ &= \text{rank}(\Sigma V^T) \\ &= \text{rank}(\Sigma) = r \end{aligned}$$

(b) Ya se sabe que $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ es un conjunto ortonormal. Por tanto, es linealmente independiente, por el Teorema 5.1. Puesto que $\mathbf{u}_i = (1/\sigma_i)A\mathbf{v}_i$ para $i = 1, \ldots, r$, cada $\mathbf{u}_i$ está en el espacio columna de A. (¿Por qué?) Más aún,



$$r = \text{rank}(A) = \dim(\text{col}(A))$$

Por tanto, $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ es una base ortonormal para col(A), por el Teorema 6.10(c).

(c) Dado que $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ es una base ortonormal para $\mathbb{R}^m$ y $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ es una base para col(A), por la propiedad (b), se tiene que $\{\mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\}$ es una base ortonormal para el complemento ortogonal de col(A). Pero $(\text{col}(A))^\perp = \text{nulo}(A^T)$ por el Teorema 5.10.

(e) Dado que

$$A\mathbf{v}_{r+1} = \cdots = A\mathbf{v}_n = \mathbf{0}$$

el conjunto $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ es un conjunto ortonormal contenido en el espacio nulo de A. Por tanto, $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ es un conjunto linealmente independiente de $n - r$ vectores en nulo(A). Pero

$$\dim(\text{nulo}(A)) = n - r$$

por el teorema del rank, de modo que $\{\mathbf{v}_{r+1}, \ldots, \mathbf{v}_n\}$ es una base ortonormal para nulo(A), por el Teorema 6.10(c).

(d) La propiedad (d) surge de la propiedad (e) y el Teorema 5.10. (En el ejercicio 32 se le pide probar esto.)

La DVS ofrece nueva comprensión geométrica acerca del efecto de las transformaciones matriciales. Muchas veces se hizo notar (sin demostración) que una matriz de $m \times n$ transforma la esfera unitaria en $\mathbb{R}^n$ en un elipsoide en $\mathbb{R}^m$. Este tema surge, por ejemplo, en las discusiones del teorema de Perron y de los operadores norma, así como en la introducción a los valores singulares en esta sección. Ahora se demostrará este resultado.

Teorema 7.16 Sea A una matriz de $m \times n$ con rank r. Entonces la imagen de la esfera unitaria en $\mathbb{R}^n$ bajo la transformación matricial que mapea $\mathbf{x}$ a $A\mathbf{x}$ es

a. la superficie de un elipsoide en $\mathbb{R}^m$ si $r = n$.
b. un elipsoide sólido en $\mathbb{R}^m$ si $r < n$.

Demostración Sea $A = U\Sigma V^T$ una descomposición de valor singular de la matriz A de $m \times n$. Sean los vectores singulares izquierdo y derecho de A $\mathbf{u}_1, \ldots, \mathbf{u}_m$ y $\mathbf{v}_1, \ldots, \mathbf{v}_n$, respectivamente. Puesto que rank$(A) = r$, los valores singulares de A satisfacen

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \qquad \text{y} \qquad \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$$

por el Teorema 7.15(a). Sea $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ un vector unitario en $\mathbb{R}^n$. Ahora, puesto que V es una matriz ortogonal, también lo es V^T, y en consecuencia $V^T\mathbf{x}$ es un vector unitario, por el Teorema 5.6. Ahora

$$V^T\mathbf{x} = \begin{bmatrix} \mathbf{v}_1^T \\ \vdots \\ \mathbf{v}_n^T \end{bmatrix}\mathbf{x} = \begin{bmatrix} \mathbf{v}_1^T\mathbf{x} \\ \vdots \\ \mathbf{v}_n^T\mathbf{x} \end{bmatrix}$$

de modo que $(\mathbf{v}_1^T\mathbf{x})^2 + \cdots + (\mathbf{v}_n^T\mathbf{x})^2 = 1$.

Por la forma de producto exterior de la DVS, se tiene $A = \sigma_1\mathbf{u}_1\mathbf{v}_1^T + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T$. Por tanto,

$$\begin{aligned} A\mathbf{x} &= \sigma_1\mathbf{u}_1\mathbf{v}_1^T\mathbf{x} + \cdots + \sigma_r\mathbf{u}_r\mathbf{v}_r^T\mathbf{x} \\ &= (\sigma_1\mathbf{v}_1^T\mathbf{x})\mathbf{u}_1 + \cdots + (\sigma_r\mathbf{v}_r^T\mathbf{x})\mathbf{u}_r \\ &= y_1\mathbf{u}_1 + \cdots + y_r\mathbf{u}_r \end{aligned}$$

donde y_i denota al escalar $\sigma_i\mathbf{v}_i^T\mathbf{x}$.

(a) Si $r = n$, entonces se debe tener $n \leq m$ y

$$\begin{aligned} A\mathbf{x} &= y_1\mathbf{u}_1 + \cdots + y_n\mathbf{u}_n \\ &= Uy \end{aligned}$$

donde $\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$. Por tanto, nuevamente por el Teorema 5.6, $\|A\mathbf{x}\| = \|U\mathbf{y}\| = \|\mathbf{y}\|$, pues U es ortogonal. Pero

$$\left(\frac{y_1}{\sigma_1}\right)^2 + \cdots + \left(\frac{y_n}{\sigma_n}\right)^2 = (\mathbf{v}_1^T\mathbf{x})^2 + \cdots + (\mathbf{v}_n^T\mathbf{x})^2 = 1$$



lo que demuestra que los vectores $A\mathbf{x}$ forman la superficie de un elipsoide en $\mathbb{R}^m$. (¿Por qué?)

(b) Si $r < n$, la única diferencia en los pasos anteriores es que la ecuación se convierte en

$$\left(\frac{y_1}{\sigma_1}\right)^2 + \cdots + \left(\frac{y_r}{\sigma_r}\right)^2 \leq 1$$

pues faltan algunos términos. Esta desigualdad corresponde a un elipsoide sólido en $\mathbb{R}^m$.

Ejemplo 7.35

Describa la imagen de la esfera unitaria en $\mathbb{R}^3$ bajo la acción de la matriz

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Solución En el ejemplo 7.34(a) se encontró la siguiente DVS de A:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix}$$

Puesto que $r = \text{rank}(A) = 2 < 3 = n$, se aplica la segunda parte del Teorema 7.16. La imagen de la esfera unitaria satisfará la desigualdad

$$\left(\frac{y_1}{\sqrt{2}}\right)^2 + \left(\frac{y_2}{1}\right)^2 \leq 1 \quad \text{o} \quad \frac{y_1^2}{2} + y_2^2 \leq 1$$

en relación con los ejes coordenados y_1y_2 en $\mathbb{R}^2$ (correspondientes a los vectores singulares izquierdos $\mathbf{u}_1$ y $\mathbf{u}_2$). Dado que $\mathbf{u}_1 = \mathbf{e}_1$ y $\mathbf{u}_2 = \mathbf{e}_2$, la imagen es como se muestra en la figura 7.20.

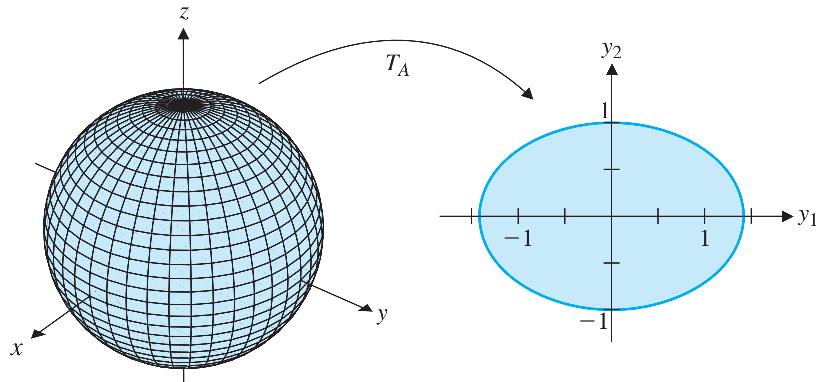
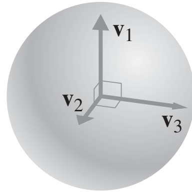


Figura 7.20

En general, es posible describir el efecto de una matriz A de $m \times n$ sobre la esfera unitaria en $\mathbb{R}^n$ en términos del efecto de cada factor en su DVS, $A = U\Sigma V^T$, de derecha a izquierda. Puesto que V^T es una matriz ortogonal, mapea la esfera unitaria en sí misma. La matriz Σ de $m \times n$ hace dos cosas: las entradas diagonales $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$ colapsan $n - r$ de las dimensiones de la esfera unitaria, lo que deja una esfera unitaria r-dimensional, que entonces las entradas diagonales distintas de cero $\sigma_1, \ldots, \sigma_r$ distorsionan en un elipsoide. Entonces la matriz ortogonal U alinea los ejes de este elipsoide con los vectores base ortonormales $\mathbf{u}_1, \ldots, \mathbf{u}_r$ en $\mathbb{R}^m$. (Vea la figura 7.21.)

Aplicaciones de la DVS

La descomposición de valor singular es una herramienta extremadamente útil, tanto práctica como teóricamente. Se observarán sólo algunas de sus muchas aplicaciones.



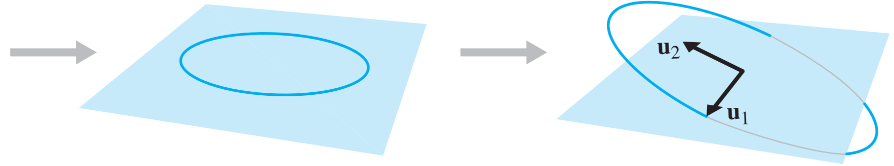


Figura 7.21

Rank Hasta el momento no ha tenido que preocuparse por calcular el rank de una matriz desde un punto de vista computacional. El rank de una matriz se calcula al reducirla por renglones a la forma escalonada y a contar el número de renglones distintos de cero. Sin embargo, como ha visto, los errores de redondeo pueden afectar este proceso, en especial si la matriz está mal condicionada. Las entradas que deben ser cero pueden terminar como números distintos de cero muy pequeños, lo que afecta la capacidad para determinar con precisión el rank y otras cantidades asociadas con la matriz. En la práctica, con frecuencia se usa la DVS para encontrar el rank de una matriz, pues es mucho más confiable cuando se presentan errores de redondeo. La idea básica detrás de este enfoque es que las matrices ortogonales U y V en la DVS preservan las longitudes y por tanto no introducen errores adicionales; cualquier error que ocurra tenderá a mostrarse en la matriz Σ.

CAS **Ejemplo 7.36**

Sea

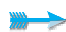
$$A = \begin{bmatrix} 8.1650 & -0.0041 & -0.0041 \\ 4.0825 & -3.9960 & 4.0042 \\ 4.0825 & 4.0042 & -3.9960 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 8.17 & 0 & 0 \\ 4.08 & -4 & 4 \\ 4.08 & 4 & -4 \end{bmatrix}$$

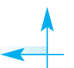
La matriz B se obtuvo mediante redondeo de las entradas en A a dos lugares decimales. Si calcula los ranks de estas dos matrices aproximadamente iguales, se encuentra que $\text{rank}(A) = 3$, pero $\text{rank}(B) = 2$. Por el teorema fundamental, esto implica, entre otras cosas, que A es invertible, pero B no lo es.

La explicación para esta diferencia crucial entre dos matrices que son aproximadamente iguales se encuentra en su DVS. Los valores singulares de A son 10, 8 y 0.01, de modo que A tiene rank 3. Los valores singulares de B son 10, 8 y 0, de modo que B tiene rank 2.

En aplicaciones prácticas, con frecuencia se supone que si un valor singular se calcula cercano a cero, entonces se arrastra error de redondeo y el valor real debe ser cero. De esta forma, puede filtrarse el "ruido". En este ejemplo, si calcula $A = U\Sigma V^T$ y sustituye

$$\Sigma = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 0.01 \end{bmatrix} \quad \text{por} \quad \Sigma' = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$


entonces $U\Sigma' V^T = B$. (¡Inténtelo!)



Normas matriciales y el número de condición La DVS puede proporcionar fórmulas simples para ciertas expresiones que involucran normas matriciales. Considere, por ejemplo, la norma de Frobenius de una matriz. El siguiente teorema demuestra que está completamente determinada por los valores singulares de la matriz.

Teorema 7.17 Sea A una matriz de $m \times n$ y sean $\sigma_1, \ldots, \sigma_r$ todos los valores singulares distintos de cero de A. Entonces

$$\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$$

La demostración de este resultado depende de la siguiente analogía del Teorema 5.6:

Si A es una matriz de $m \times n$ y Q es una matriz ortogonal de $m \times m$, entonces

$$\|QA\|_F = \|A\|_F \qquad (2)$$

Para demostrar que esto es cierto, calcule

$$\begin{aligned}\|QA\|_F^2 &= \|[Q\mathbf{a}_1 \quad \cdots \quad Q\mathbf{a}_n]\|_F^2 \\ &= \|Q\mathbf{a}_1\|_E^2 + \cdots + \|Q\mathbf{a}_n\|_E^2 \\ &= \|\mathbf{a}_1\|_E^2 + \cdots + \|\mathbf{a}_n\|_E^2 \\ &= \|A\|_F^2\end{aligned}$$

Demostración del Teorema 7.17 Sea $A = U\Sigma V^T$ una descomposición de valor singular de A. Entonces, al usar la ecuación (2) dos veces, se tiene

$$\begin{aligned}\|A\|_F^2 &= \|U\Sigma V^T\|_F^2 \\ &= \|\Sigma V^T\|_F^2 = \|(\Sigma V^T)^T\|_F^2 \\ &= \|V\Sigma^T\|_F^2 = \|\Sigma^T\|_F^2 = \sigma_1^2 + \cdots + \sigma_r^2\end{aligned}$$

que establece el resultado.

CAS **Ejemplo 7.37** Verifique el Teorema 7.17 para la matriz A en el ejemplo 7.18.

Solución La matriz $A = \begin{bmatrix} 3 & -1 \\ 2 & 4 \end{bmatrix}$ tiene valores singulares 4.5150 y 3.1008. Se comprueba que

$$\sqrt{4.5150^2 + 3.1008^2} = \sqrt{30} = \|A\|_F$$

que concuerda con el ejemplo 7.18.

En la sección 7.2 se comentó que no hay una fórmula sencilla para el operador 2-norma de una matriz A. Aunque esto es cierto, la DVS de A proporciona una expresión muy amigable para $\|A\|_2$. Recuerde que

$$\|A\|_2 = \max_{\|\mathbf{x}\|=1} \|A\mathbf{x}\|$$

donde la norma vectorial es la norma euclidiana ordinaria. Por el Teorema 7.16, para $\|\mathbf{x}\| = 1$, el conjunto de vectores $\|A\mathbf{x}\|$ yace en o dentro de un elipsoide cuyos semiejes

tienen longitudes iguales a los valores singulares de A distintos de cero. Se tiene inmediatamente que el más grande de ellos es σ_1, de modo que

$$\|A\|_2 = \sigma_1$$

Esto proporciona una manera elegante de expresar el número de condición de una matriz (cuadrada) con respeto al operador 2-norma. Recuerde que el número de condición (con respecto al operador 2-norma) de una matriz invertible A se define como

$$\text{cond}_2(A) = \|A^{-1}\|_2\|A\|_2$$

Como se le pedirá demostrar en el ejercicio 28, si $A = U\Sigma V^T$, entonces $A^{-1} = V\Sigma^{-1}U^T$. Por tanto, los valores singulares de A^{-1} son $1/\sigma_1, \ldots, 1/\sigma_n$ (¿por qué?) y

$$1/\sigma_n \geq \cdots \geq 1/\sigma_1$$

Se tiene que $\|A^{-1}\|_2 = 1/\sigma_n$, de modo que

$$\text{cond}_2(A) = \frac{\sigma_1}{\sigma_n}$$

Ejemplo 7.38

Encuentre el número de condición 2 de la matriz A en el ejemplo 7.36.

Solución Dado que $\sigma_1 = 10$ y $\sigma_3 = 0.01$,

$$\text{cond}_2(A) = \frac{\sigma_1}{\sigma_3} = \frac{10}{0.01} = 1000$$

Este valor es suficientemente grande para sugerir que A puede estar mal condicionada y debe estar alerta del efecto de los errores de redondeo.

La pseudoinversa y la aproximación por mínimos cuadrados En la sección 7.3 se produjo la fórmula $A^+ = (A^TA)^{-1}A^T$ para la pseudoinversa de una matriz A. Claramente, esta fórmula sólo es válida si A^TA es invertible, como se hizo notar en su momento. Equipado con la DVS, ahora puede definir la pseudoinversa de *cualquier* matriz y generalizar la fórmula anterior.

E. H. Moore (1862-1932) fue un matemático estadounidense que trabajó en teoría de grupos, teoría de números y geometría. Fue el primer director del departamento de matemáticas de la Universidad de Chicago cuando abrió en 1892. En 1920 introdujo una matriz generalizada inversa que incluyó matrices rectangulares. Su obra no recibió mucha atención debido a su oscuro estilo de escritura.

Definición Sea $A = U\Sigma V^T$ una DVS para una matriz A de $m \times n$, donde $\Sigma = \begin{bmatrix} D & O \\ O & O \end{bmatrix}$ y D es una matriz diagonal de $r \times r$ que contiene los valores singulares distintos de cero $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ de A. La ***pseudoinversa*** (o ***inversa de Moore-Penrose***) de A es la matriz A^+ de $n \times m$ definida por

$$A^+ = V\Sigma^+U^T$$

donde Σ^+ es la matriz de $n \times m$

$$\Sigma^+ = \begin{bmatrix} D^{-1} & O \\ O & O \end{bmatrix}$$

Ejemplo 7.39

Encuentre las pseudoinversas de las matrices del ejemplo 7.34.

Solución (a) A partir de la DVS

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} = U\Sigma V^T$$

se forma

$$\Sigma^+ = \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Entonces

$$A^+ = V\Sigma^+U^T = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 1/2 & 0 \\ 0 & 1 \end{bmatrix}$$

(b) Se tiene la DVS

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = U\Sigma V^T$$

de modo que

$$\Sigma^+ = \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

y

$$A^+ = V\Sigma^+U^T = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1/3 & 2/3 & -1/3 \\ 1/3 & -1/3 & 2/3 \end{bmatrix}$$

Es directo comprobar que esta nueva definición de la pseudoinversa generaliza la anterior, pues si la matriz de $A = U\Sigma V^T$ $m \times n$ tiene columnas linealmente independientes, entonces la sustitución directa muestra que $(A^TA)^{-1}A^T = V\Sigma^+U^T$. (En el ejercicio 50 se le pide verificar esto.) En los ejercicios se exploran otras propiedades de la pseudoinversa.

Se vio que cuando A tiene columnas linealmente independientes, existe una solución única $\overline{\mathbf{x}}$ de mínimos cuadrados para $A\mathbf{x} = \mathbf{b}$; esto es, las ecuaciones normales $A^TA\mathbf{x} = A^T\mathbf{b}$ tienen la solución única

$$\overline{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b} = A^+\mathbf{b}$$

Cuando las columnas de A son linealmente dependientes, entonces A^TA no es invertible, de modo que las ecuaciones normales tienen infinitas soluciones. En este caso, se pedirá la solución $\overline{\mathbf{x}}$ de *longitud mínima* (es decir, la más cercana al origen). Es evidente que esta vez simplemente se usa la versión general de la pseudoinversa.

Uno de los que estuvo al tanto del trabajo de Moore acerca de los inversos matriciales fue Roger Penrose (n. 1931), quien introdujo su propia noción de un inverso matricial generalizado en 1955. Penrose hizo muchas aportaciones a la geometría y la física teórica. También es el inventor de un tipo de *mosaico no periódico* que cubre el plano sólo con dos formas diferentes de mosaico, pero no tiene patrón. Recibió muchos premios, incluido el Premio Wolf de Física en 1988, que compartió con Stephen Hawking. En 1994 fue nombrado caballero por sus servicios a la ciencia. Sir Roger Penrose es en la actualidad profesor emérito Rouse Ball de matemáticas en la Universidad de Oxford.

Teorema 7.18 El problema de mínimos cuadrados $A\mathbf{x} = \mathbf{b}$ tiene una solución única $\overline{\mathbf{x}}$ de mínimos cuadrados de longitud mínima que está dada por

$$\overline{\mathbf{x}} = A^{+}\mathbf{b}$$

Demostración Sea A una matriz de $m \times n$ de rank r con DVS: $A = U\Sigma V^T$ (de modo que $A^+ = V\Sigma^+ U^T$). Sean $\mathbf{y} = V^T\mathbf{x}$ y $\mathbf{c} = U^T\mathbf{b}$. Escriba $\mathbf{y}$ y $\mathbf{c}$ en forma de bloque como

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \quad \text{y} \quad \mathbf{c} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix}$$

donde $\mathbf{y}_1$ y $\mathbf{c}_1$ están en $\mathbb{R}^r$.

Se quiere minimizar $\|\mathbf{b} - A\mathbf{x}\|$ o, de manera equivalente, $\|\mathbf{b} - A\mathbf{x}\|^2$. Al usar el Teorema 5.6 y el hecho de que U^T es ortogonal (porque U lo es), se tiene

$$\begin{aligned} \|\mathbf{b} - A\mathbf{x}\|^2 &= \|U^T(\mathbf{b} - A\mathbf{x})\|^2 = \|U^T(\mathbf{b} - U\Sigma V^T\mathbf{x})\|^2 = \|U^T\mathbf{b} - U^TU\Sigma V^T\mathbf{x}\|^2 \\ &= \|\mathbf{c} - \Sigma\mathbf{y}\|^2 = \left\| \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} - \begin{bmatrix} D & O \\ O & O \end{bmatrix}\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \mathbf{c}_1 - D\mathbf{y}_1 \\ \mathbf{c}_2 \end{bmatrix} \right\|^2 \end{aligned}$$

La única parte de esta expresión sobre la que se tiene control es $\mathbf{y}_1$, de modo que el valor mínimo ocurre cuando $\mathbf{c}_1 - D\mathbf{y}_1 = \mathbf{0}$ o, de manera equivalente, cuando $\mathbf{y}_1 = D^{-1}\mathbf{c}_1$. De modo que todas las soluciones $\mathbf{x}$ de mínimos cuadrados son de la forma

$$\mathbf{x} = V\mathbf{y} = V\begin{bmatrix} D^{-1}\mathbf{c}_1 \\ \mathbf{y}_2 \end{bmatrix}$$

Establezca

$$\overline{\mathbf{x}} = V\overline{\mathbf{y}} = V\begin{bmatrix} D^{-1}\mathbf{c}_1 \\ \mathbf{0} \end{bmatrix}$$

Se afirma que esta $\overline{\mathbf{x}}$ es la solución de mínimos cuadrados de longitud mínima. Para demostrar esto, suponga que

$$\mathbf{x}' = V\mathbf{y}' = V\begin{bmatrix} D^{-1}\mathbf{c}_1 \\ \mathbf{y}_2 \end{bmatrix}$$

es una solución diferente de mínimos cuadrados (por tanto, $\mathbf{y}_2 \neq \mathbf{0}$). Entonces

$$\|\overline{\mathbf{x}}\| = \|V\overline{\mathbf{y}}\| = \|\overline{\mathbf{y}}\| < \|\mathbf{y}'\| = \|V\mathbf{y}'\| = \|\mathbf{x}'\|$$

como se afirmó.

Todavía debe demostrar que $\overline{\mathbf{x}}$ es igual a $A^+\mathbf{b}$. Para hacerlo, simplemente calcule

$$\begin{aligned} \overline{\mathbf{x}} = V\overline{\mathbf{y}} &= V\begin{bmatrix} D^{-1}\mathbf{c}_1 \\ \mathbf{0} \end{bmatrix} = V\begin{bmatrix} D^{-1} & O \\ O & O \end{bmatrix}\begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} \\ &= V\Sigma^+\mathbf{c} = V\Sigma^+U^T\mathbf{b} = A^+\mathbf{b} \end{aligned}$$

Ejemplo 7.40 Encuentre la solución de mínimos cuadrados de longitud mínima para $A\mathbf{x} = \mathbf{b}$, donde

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{y} \quad \mathbf{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Solución Las ecuaciones correspondientes

$$x + y = 0$$
$$x + y = 1$$

claramente son inconsistentes, de modo que la solución de mínimos cuadrados es la única esperanza. Más aún, las columnas de A son linealmente dependientes, de modo que habrá infinitas soluciones de mínimos cuadrados, entre las cuales se quiere aquella con longitud mínima.

Una DVS de A está dada por

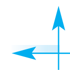
$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = U\Sigma V^T$$



(Verifique esto.) Se tiene que

$$A^+ = V\Sigma^+U^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/4 & 1/4 \\ 1/4 & 1/4 \end{bmatrix}$$

de modo que
$$\overline{\mathbf{x}} = A^+\mathbf{b} = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}$$



Puede ver que la solución mínima de mínimos cuadrados del ejemplo 7.40 satisface $x + y = \frac{1}{2}$. En cierto sentido, este es un compromiso entre las dos ecuaciones con las que se comenzó. En el ejercicio 49 se le pide resolver directamente las ecuaciones normales para este problema y verificar que esta solución realmente es la más cercana al origen.

El teorema fundamental de las matrices invertibles Es adecuado concluir con la revisión, una vez más, del teorema fundamental de las matrices invertibles. No es de sorprender que los valores singulares de una matriz cuadrada digan cuándo la matriz es invertible.

Teorema 7.19 **El teorema fundamental de las matrices invertibles: versión final**

Sea A una matriz de $n \times n$ y sea $T : V \rightarrow W$ una transformación lineal cuya matriz $[T]_{\mathcal{C}\leftarrow\mathcal{B}}$ con respecto a las bases $\mathcal{B}$ y $\mathcal{C}$ de V y W, respectivamente, es A. Los siguientes enunciados son equivalentes:

a. A es invertible.
b. $A\mathbf{x} = \mathbf{b}$ tiene una solución única para toda $\mathbf{b}$ en $\mathbb{R}^n$
c. $A\mathbf{x} = \mathbf{0}$ sólo tiene la solución trivial.
d. La forma escalonada reducida por renglón de A es I_n.
e. A es un producto de matrices elementales.
f. $\text{rank}(A) = n$
g. $\text{nulidad}(A) = 0$
h. Los vectores columna de A son linealmente independientes.
i. Los vectores columna de A generan a $\mathbb{R}^n$.
j. Los vectores columna de A forman una base para $\mathbb{R}^n$.
k. Los vectores renglón de A son linealmente independientes.
l. Los vectores renglón de A generan a $\mathbb{R}^n$.

m. Los vectores renglón de A forman una base para $\mathbb{R}^n$.
n. $\det A \neq 0$
o. 0 no es un eigenvalor de A.
p. T es invertible.
q. T es uno a uno.
r. T es sobreyectiva.
s. $\ker(T) = \{\mathbf{0}\}$
t. Rango $(T) = W$
u. 0 no es un valor singular de A.

Demostración Note primero que, por la definición de valores singulares, 0 es un valor singular de A si y sólo si 0 es un eigenvalor de A^TA.

(a) $\Rightarrow$ (u) Si A es invertible, también lo es A^T, y por tanto también lo es A^TA. Por tanto, la propiedad (o) implica que 0 no es un eigenvalor de A^TA, de modo que 0 no es un valor singular de A.

(u) $\Rightarrow$ (a) Si 0 no es un valor singular de A, entonces 0 no es un eigenvalor de A^TA. En consecuencia, A^TA es invertible, por la equivalencia de las propiedades (a) y (o). Pero entonces rank$(A) = n$, por el Teorema 3.28, de modo que A es invertible, por la equivalencia de las propiedades (a) y (f).

Compresión de imágenes digitales

Entre las muchas aplicaciones de la DVS, una de las más impresionantes es su uso en la compresión de imágenes digitales, de modo que puedan transmitirse electrónicamente con eficiencia (por satélite, fax, Internet o similares). Ya se analizó el problema de detectar y corregir errores en tales transmisiones. El problema que ahora se quiere considerar tiene que ver con la reducción de la cantidad de información que debe transmitirse sin perder información esencial.

En el caso de las imágenes digitales, suponga que tiene una imagen en escala de grises con un tamaño de 340×280 píxeles. Cada píxel es una de los 256 tonos de gris, que pueden representarse mediante un número entre 0 y 255. Esta información se puede almacenar en una matriz A 340×280, pero transmitir y manipular estos 95,200 números es muy costoso. La idea detrás de la compresión de imágenes es que algunas partes de la imagen son menos interesantes que otras. Por ejemplo, en una fotografía de alguien de pie en exteriores, puede haber mucho cielo en el fondo, mientras que el rostro de la persona contiene muchos detalles. Probablemente podría transmitir cada segundo o tercer píxel en el fondo, pero le gustaría mantener todos los píxeles en la región del rostro.

Es evidente que los pequeños valores singulares en la DVS de la matriz A provienen de las partes "aburridas" de la imagen, y puede ignorar muchas de ellas. Suponga, entonces, que tiene la DVS de A en forma de producto externo

$$A = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T$$

Sea $k \leq r$ y defina

$$A_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^T$$

Entonces A_k es una aproximación de A que corresponde a mantener sólo los primeros k valores singulares y los correspondientes vectores singulares. Para el ejemplo de 340×280, puede descubrir que es suficiente con transmitir sólo los datos correspondientes a los primeros 20 valores singulares. Entonces, en lugar de transmitir 95,200 números, sólo es necesario enviar 20 valores singulares más los 20 vectores $\mathbf{u}_1, \ldots, \mathbf{u}_{20}$ en $\mathbb{R}^{340}$ y los 20 vectores $\mathbf{v}_1, \ldots, \mathbf{v}_{20}$ en $\mathbb{R}^{280}$, para un total de

$$20 + 20 \cdot 340 + 20 \cdot 280 = 12{,}420$$

números. ¡Esto representa un ahorro sustancial!


© Bettmann/CORBIS

Figura 7.22

La imagen del matemático Gauss en la figura 7.22 es una imagen de 340×280 píxeles. Tiene 256 tonos de gris, de modo que la matriz correspondiente A tiene 340×280, con entradas entre 0 y 255.

Es evidente que la matriz A tiene rank 280. Si aproxima A mediante A_k, como se describió líneas arriba, se obtiene una imagen que corresponde a los primeros k valores singulares de A. La figura 7.23 muestra varias de estas imágenes para valores de k de 2 a 256. Al principio, la imagen es muy borrosa, pero muy rápidamente adquiere forma. Note que A_{32} ya produce una aproximación bastante buena de la imagen real (que proviene de $A = A_{280}$, como se muestra en la esquina superior izquierda de la figura 7.23).

Algunos de los valores singulares de A son $\sigma_1 = 49{,}096$, $\sigma_{16} = 22{,}589$, $\sigma_{32} = 10{,}187$, $\sigma_{64} = 484$, $\sigma_{128} = 182$, $\sigma_{256} = 5$ y $\sigma_{280} = 0.5$. Los valores singulares más pequeños contribuyen muy poco a la imagen, razón por la cual las aproximaciones rápidamente parecen tan cercanas al original.

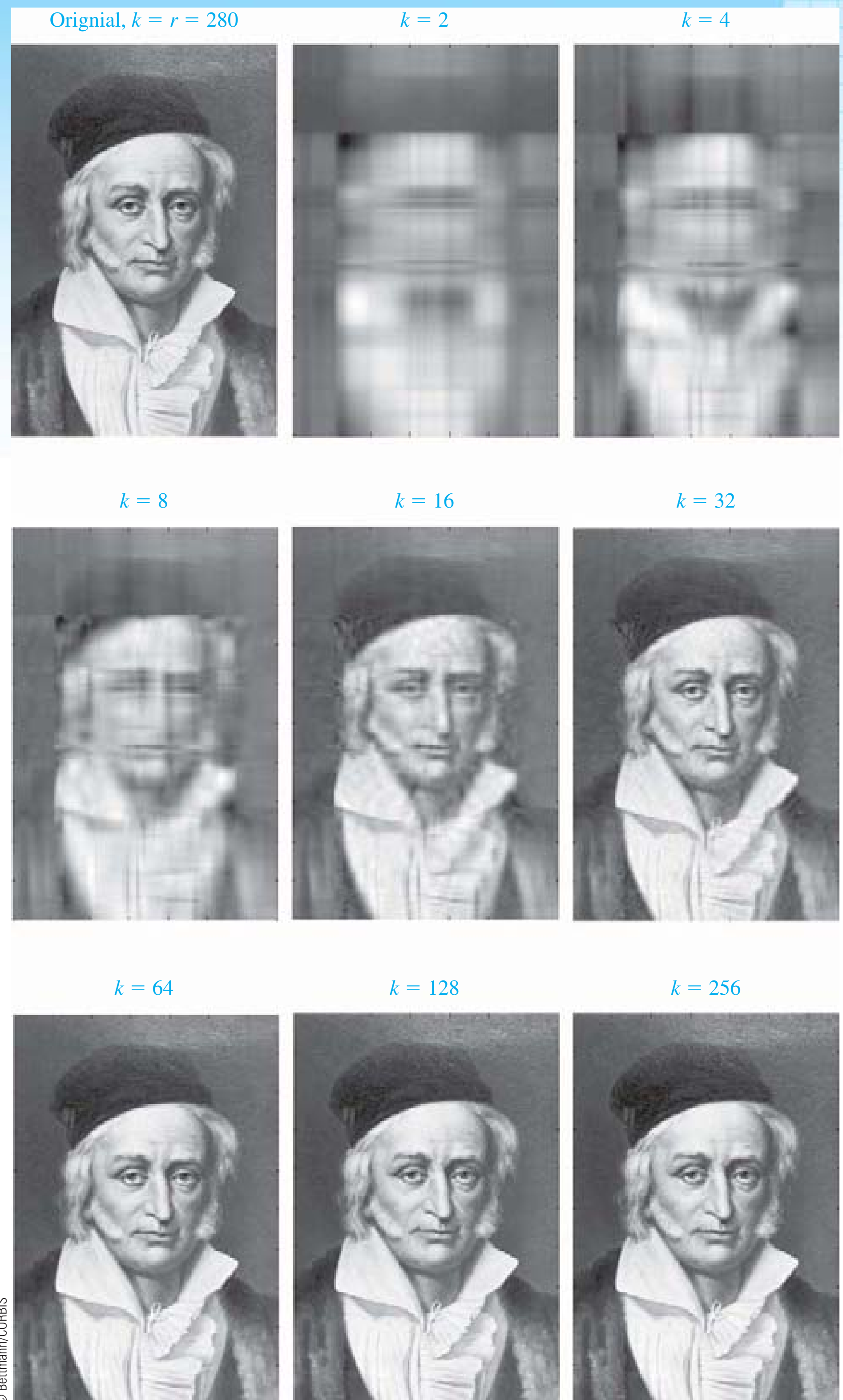


Figura 7.23

Ejercicios 7.4

En los ejercicios 1-10, encuentre los valores singulares de la matriz dada.

1. $A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ **2.** $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ **4.** $A = \begin{bmatrix} \sqrt{2} & 1 \\ 0 & \sqrt{2} \end{bmatrix}$

5. $A = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ **6.** $A = \begin{bmatrix} 3 & 4 \end{bmatrix}$

7. $A = \begin{bmatrix} 0 & 0 \\ 0 & 3 \\ -2 & 0 \end{bmatrix}$ **8.** $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -2 & 2 \end{bmatrix}$

9. $A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \end{bmatrix}$ **10.** $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -3 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

En los ejercicios 11-20, encuentre una DVS de la matriz indicada.

11. A en el ejercicio 3 **12.** $A = \begin{bmatrix} -2 & 0 \\ 0 & 0 \end{bmatrix}$

13. $A = \begin{bmatrix} 0 & -2 \\ -3 & 0 \end{bmatrix}$ **14.** $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

15. A en el ejercicio 5 **16.** A en el ejercicio 6

17. A en el ejercicio 7 **18.** A en el ejercicio 8

19. A en el ejercicio 9 **20.** $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$

En los ejercicios 21-24, encuentre la forma de producto externo de la DVS para la matriz en los ejercicios dados.

21. Ejercicios 3 y 11 **22.** Ejercicios 14

23. Ejercicios 7 y 17 **24.** Ejercicios 9 y 19

25. Demuestre que las matrices U y V en la DVS no están determinadas de manera única. [*Sugerencia:* encuentre un ejemplo en el que sería posible hacer diferentes elecciones en la construcción de dichas matrices.]

26. Sea A una matriz simétrica. Demuestre que los valores singulares de A son:

(a) los valores absolutos de los eigenvalores de A.

(b) los eigenvalores de A si A es definida positiva.

27. **(a)** Demuestre que para una matriz simétrica A definida positiva, el Teorema 7.13 produce la diagonalización ortogonal de A, como garantiza el teorema espectral.

(b) Demuestre que, para una matriz simétrica A definida positiva, el Teorema 7.14 produce la descomposición espectral de A.

28. Si A es una matriz invertible con DVS: $A = U\Sigma V^T$, demuestre que Σ es invertible y que $A^{-1} = V\Sigma^{-1}U^T$ es una DVS de A^{-1}.

29. Demuestre que si $A = U\Sigma V^T$ es una DVS de A, entonces los vectores singulares izquierdos son eigenvectores de AA^T.

30. Demuestre que A y A^T tienen los mismos valores singulares.

31. Sea Q una matriz ortogonal tal que QA tiene sentido. Demuestre que A y QA tienen los mismos valores singulares.

32. Demuestre el Teorema 7.15(d).

33. ¿Cuál es la imagen del círculo unitario en $\mathbb{R}^3$ bajo la acción de la matriz en el ejercicio 3?

34. ¿Cuál es la imagen del círculo unitario en $\mathbb{R}^3$ bajo la acción de la matriz en el ejercicio 7?

35. ¿Cuál es la imagen de la esfera unitaria en $\mathbb{R}^3$ bajo la acción de la matriz en el ejercicio 9?

36. ¿Cuál es la imagen de la esfera unitaria en $\mathbb{R}^3$ bajo la acción de la matriz en el ejercicio 10?

En los ejercicios 37-40, calcule (a) $\|A\|_2$ y (b) $\mathrm{cond}_2(A)$ para la matriz indicada.

37. A en el ejercicio 3 **38.** A en el ejercicio 8

39. $A = \begin{bmatrix} 1 & 0.9 \\ 1 & 1 \end{bmatrix}$ **40.** $A = \begin{bmatrix} 10 & 10 & 0 \\ 100 & 100 & 1 \end{bmatrix}$

En los ejercicios 41-44, calcule la pseudoinversa A^+ de A en el ejercicio dado.

41. Ejercicio 3 **42.** Ejercicio 8

43. Ejercicio 9 **44.** Ejercicio 10

En los ejercicios 45-48, encuentre A^+ y úsela para calcular la solución de mínimos cuadrados de longitud mínima para $A\mathbf{x} = \mathbf{b}$.

45. $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}$

46. $A = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$

47. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

48. $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

49. **(a)** Establezca y resuelva las ecuaciones normales para el sistema de ecuaciones en el ejemplo 7.40.
(b) Encuentre una expresión paramétrica para la longitud de un vector solución en el inciso (a).
(c) Encuentre el vector solución de longitud mínima y verifique que es el producido por el método del ejemplo 7.40. [*Sugerencia:* recuerde cómo encontrar las coordenadas del vértice de una parábola.]

50. Verifique que cuando A tiene columnas linealmente independientes, las definiciones de pseudoinversa en esta sección y en la sección 7.3 son iguales.

51. Verifique que la pseudoinversa (como se definió en esta sección) satisface las condiciones de Penrose para A (Teorema 7.12 en la sección 7.3).

52. Demuestre que A^+ es la *única* matriz que satisface las condiciones de Penrose para A. Para hacer esto, suponga que A' es una matriz que satisface las condiciones de Penrose: (a) $AA'A = A$, (b) $A'AA' = A'$, y (c) AA' y $A'A$ son simétricas. Pruebe que $A' = A^+$. [*Sugerencia:* use las condiciones de Penrose para A^+ y A' para demostrar que $A^+ = A'AA^+$ y $A' = A'AA^+$. Es útil notar que la condición (c) puede escribirse como $AA' = (A')^TA^T$ y $A'A = A^T(A')^T$, con versiones similares para A^+.]

53. Demuestre que $(A^+)^+ = A$. [*Sugerencia:* demuestre que A satisface las condiciones de Penrose para A^+. Por tanto, por el ejercicio 52, A debe ser $(A^+)^+$.]

54. Demuestre que $(A^+)^T = (A^T)^+$. [*Sugerencia:* demuestre que $(A^+)^T$ satisface las condiciones de Penrose para A^T. Por tanto, por el ejercicio 52, $(A^+)^T$ debe ser $(A^T)^+$.]

55. Demuestre que si A es una matriz simétrica idempotente, entonces $A^+ = A$.

56. Sea Q una matriz ortogonal tal que QA tiene sentido. Demuestre que $(QA)^+ = A^+Q^T$.

57. Demuestre que si A es una matriz definida positiva con DVS: $A = U\Sigma V^T$, entonces $U = V$.

58. Demuestre que para una matriz diagonal, las 1-, 2- y ∞-normas son iguales.

59. Demuestre que para cualquier matriz cuadrada A, $\|A\|_2^2 \le \|A\|_1\|A\|_\infty$. [*Sugerencia:* $\|A\|_2^2$ es el cuadrado del valor singular más grande de A y por tanto es igual al eigenvalor más grande de A^TA. Ahora use el ejercicio 34 de la sección 7.2.]

a + bi *Todo número complejo puede escribirse en forma polar como $z = re^{i\theta}$, donde $r = |z|$ es un número real no negativo y θ es su argumento, con $|e^{i\theta}| = 1$. (Vea el Apéndice C.) Por tanto, z se descompuso en un factor de elongación r y un factor de rotación $e^{i\theta}$. Existe una descomposición análoga $A = RQ$ para matrices cuadradas llamada* ***descomposición polar***.

60. Demuestre que toda matriz cuadrada A puede factorizarse como $A = RQ$, donde R es simétrica semidefinida positiva y Q es ortogonal. [*Sugerencia:* demuestre que la DVS puede reescribirse para producir

$$A = U\Sigma V^T = U\Sigma(U^TU)V^T = (U\Sigma U^T)(UV^T)$$

Luego demuestre que $R = U\Sigma U^T$ y $Q = UV^T$ tienen las propiedades correctas.]

Encuentre una descomposición polar de las matrices en los ejercicios 61-64.

61. A en el ejercicio 3

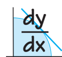
62. A en el ejercicio 14

63. $A = \begin{bmatrix} 1 & 2 \\ -3 & -1 \end{bmatrix}$

64. $A = \begin{bmatrix} 4 & 2 & -3 \\ -2 & 2 & 6 \\ 4 & -1 & 6 \end{bmatrix}$

7.5 Aplicaciones

Aproximación de funciones



En muchas aplicaciones es necesario aproximar una función dada mediante una función "más amigable". Por ejemplo, acaso quiera aproximar $f(x) = e^x$ mediante una función lineal $g(x) = c + dx$ en cierto intervalo $[a, b]$. En este caso, se tiene una función continua f y se quiere aproximarla tan cercanamente como sea posible en el intervalo $[a, b]$

mediante una función g en el subespacio $\mathscr{P}_1$. El problema general puede parafrasearse del modo siguiente:

Dada una función continua f sobre un intervalo $[a, b]$ y un subespacio W de $\mathscr{C}[a, b]$, encuentre la función "más cercana" a f en W.

El problema es análogo al ajuste de puntos de datos mediante mínimos cuadrados , excepto que ahora se tienen infinitos puntos de datos, a saber: los puntos sobre la gráfica de la función f. ¿Qué significaría "aproximar" en este contexto? Una vez más, el teorema de mejor aproximación tiene la respuesta.

La función dada f existe en el espacio vectorial $\mathscr{C}[a, b]$ de funciones continuas sobre el intervalo $[a, b]$. Este es un espacio con producto interno, con el siguiente producto interno

$$\langle f, g\rangle = \int_a^b f(x)g(x)\,dx$$

Si W es un subespacio con dimensión finita de $\mathscr{C}[a, b]$, entonces la mejor aproximación a f en W está dada por la proyección de f sobre W, por el Teorema 7.8. Más aún, si $\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ es una base ortogonal para W, entonces

$$\text{proy}_W(f) = \frac{\langle \mathbf{u}_1, f\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}\mathbf{u}_1 + \cdots + \frac{\langle \mathbf{u}_k, f\rangle}{\langle \mathbf{u}_k, \mathbf{u}_k\rangle}\mathbf{u}_k$$

Ejemplo 7.41

Encuentre la mejor aproximación lineal a $f(x) = e^x$ sobre el intervalo $[-1, 1]$.

Solución Las funciones lineales son polinomios de grado 1, de modo que se usa el subespacio $W = \mathscr{P}_1[-1, 1]$ de $\mathscr{C}[-1, 1]$ con el producto interno

$$\langle f, g\rangle = \int_{-1}^1 f(x)g(x)\,dx$$

Una base para $\mathscr{P}_1[-1, 1]$ está dada por $\{1, x\}$. Dado que

$$\langle 1, x\rangle = \int_{-1}^1 x\,dx = 0$$

ésta es una base ortogonal, de modo que la mejor aproximación para f en W es

$$\begin{aligned} g(x) = \text{proy}_W(e^x) &= \frac{\langle 1, e^x\rangle}{\langle 1, 1\rangle}1 + \frac{\langle x, e^x\rangle}{\langle x, x\rangle}x \\ &= \frac{\int_{-1}^1 (1 \cdot e^x)\,dx}{\int_{-1}^1 (1 \cdot 1)\,dx} + \frac{\int_{-1}^1 xe^x\,dx}{\int_{-1}^1 x^2\,dx}x \\ &= \frac{e - e^{-1}}{2} + \frac{2e^{-1}}{\frac{2}{3}}x \\ &= \tfrac{1}{2}(e - e^{-1}) + 3e^{-1}x \approx 1.18 + 1.10x \end{aligned}$$

donde se usó integración por partes para evaluar $\int_{-1}^{1} xe^x\, dx$. (Compruebe estos cálculos.) Vea la figura 7.24.

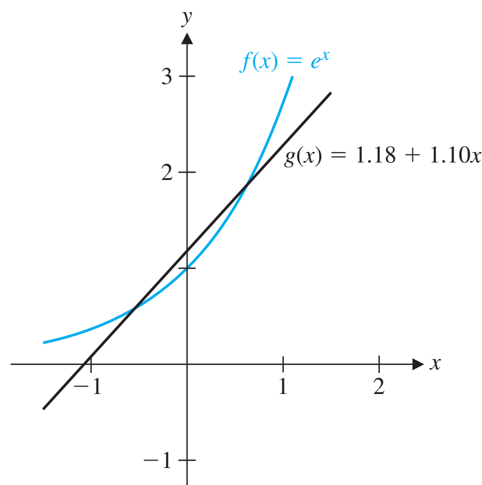


Figura 7.24

El error al aproximar f mediante g está especificado por el teorema de mejor aproximación: la distancia $\|f - g\|$ entre f y g en relación con el producto interno sobre $\mathscr{C}[-1, 1]$. Este error es justo

$$\|f - g\| = \sqrt{\int_{-1}^{1} (f(x) - g(x))^2\, dx}$$

y con frecuencia se llama ***error cuadrático medio***. Con la ayuda de un CAS, se descubre que el error cuadrático medio en el ejemplo 7.41 es

$$\|e^x - (\tfrac{1}{2}(e - e^{-1}) + 3e^{-1}x)\| = \sqrt{\int_{-1}^{1} (e^x - \tfrac{1}{2}(e - e^{-1}) - 3e^{-1}x)^2\, dx} \approx 0.23$$

Comentario El error cuadrático medio puede considerarse como el análogo al área entre las gráficas de f y g sobre el intervalo especificado. Recuerde que el área entre las gráficas de f y g sobre el intervalo $[a, b]$ está dada por

$$\int_a^b |f(x) - g(x)|\, dx$$

(Vea la figura 7.25.)

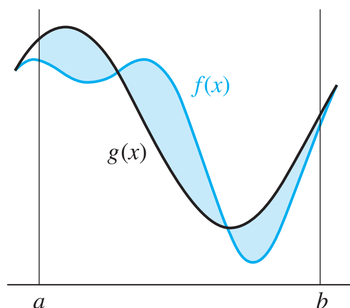


Figura 7.25

Aunque la ecuación en el comentario anterior es una medida sensible del "error" entre f y g, el signo de valor absoluto hace difícil trabajar con él. El error cuadrático medio es más fácil de usar y por tanto preferible. La raíz cuadrada es necesaria para "compensar" la elevación al cuadrado y mantener la unidad de medición igual como sería para el área entre las curvas. Para propósitos de comparación, el área entre las gráficas de f y g en el ejemplo 7.41 es

$$\int_{-1}^{1} |e^x - \tfrac{1}{2}(e - e^{-1}) - 3e^{-1}x|\, dx \approx 0.28$$

Ejemplo 7.42

Encuentre la mejor aproximación cuadrática a $f(x) = e^x$ sobre el intervalo $[-1, 1]$.

Solución Una función cuadrática es un polinomio de la forma $g(x) = a + bx + cx^2$ en $W = \mathscr{P}_2[-1, 1]$. Esta vez, la base estándar $\{1, x, x^2\}$ no es ortogonal. Sin embargo, puede construir una base ortogonal usando el proceso de Gram-Schmidt, como se hizo en el ejemplo 7.8. El resultado es el conjunto de polinomios de Legendre

$$\{1, x, x^2 - \tfrac{1}{3}\}$$

Al usar este conjunto como la base, calcule la mejor aproximación a f en W como $g(x) = \text{proy}_W(e^x)$. Los términos lineales en este cálculo son exactamente como en el ejemplo 7.41, de modo que sólo se requieren los cálculos adicionales

$$\langle x^2 - \tfrac{1}{3}, e^x \rangle = \int_{-1}^{1} (x^2 - \tfrac{1}{3})e^x\,dx = \int_{-1}^{1} x^2 e^x\,dx - \tfrac{1}{3}\int_{-1}^{1} e^x\,dx = \tfrac{2}{3}(e - 7e^{-1})$$

y $$\langle x^2 - \tfrac{1}{3}, x^2 - \tfrac{1}{3} \rangle = \int_{-1}^{1} (x^2 - \tfrac{1}{3})^2\,dx = \int_{-1}^{1} (x^4 - \tfrac{2}{3}x^2 + \tfrac{1}{9})\,dx = \tfrac{8}{45}$$

Entonces la mejor aproximación cuadrática a $f(x) = e^x$ sobre el intervalo $[-1, 1]$ es

$$\begin{aligned} g(x) = \text{proy}_W(e^x) &= \frac{\langle 1, e^x \rangle}{\langle 1, 1 \rangle}1 + \frac{\langle x, e^x \rangle}{\langle x, x \rangle}x + \frac{\langle x^2 - \frac{1}{3}, e^x \rangle}{\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3} \rangle}(x^2 - \tfrac{1}{3}) \\ &= \tfrac{1}{2}(e - e^{-1}) + 3e^{-1}x + \frac{\frac{2}{3}(e - 7e^{-1})}{\frac{8}{45}}(x^2 - \tfrac{1}{3}) \\ &= \frac{3(11e^{-1} - e)}{4} + 3e^{-1}x + \frac{15(e - 7e^{-1})}{4}x^2 \approx 1.00 + 1.10x + 0.54x^2 \end{aligned}$$

(Vea la figura 7.26.)

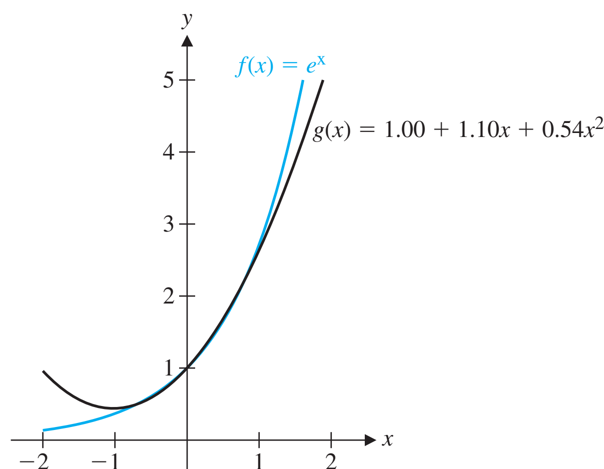


Figura 7.26

Note cuánto mejor es la aproximación cuadrática del ejemplo 7.42 que la aproximación lineal del ejemplo 7.41. Es evidente que, en el caso cuadrático, el error cuadrático medio es

$$\|e^x - g(x)\| = \sqrt{\int_{-1}^{1} (e^x - g(x))^2 \, dx} \approx 0.04$$

En general, mientras más grande sea el grado del polinomio de aproximación, menor será el error y mejor la aproximación.

En muchas aplicaciones, las funciones se aproximan mediante combinaciones de funciones seno y coseno. Este método es particularmente útil si la función por aproximar muestra comportamiento periódico o casi periódico (como el de una onda sonora, un impulso eléctrico o el movimiento de un sistema en vibración). Una función de la forma

$$\begin{aligned} p(x) = a_0 + a_1 \cos x + a_2 \cos 2x + \cdots + a_n \cos nx + b_1 \operatorname{sen} x \\ + b_2 \operatorname{sen} 2x + \cdots + b_n \operatorname{sen} nx \end{aligned} \quad (1)$$

se llama ***polinomio trigonométrico;*** si a_n y b_n no son ambos cero, entonces se dice que $p(x)$ tiene ***orden n***. Por ejemplo,

$$p(x) = 3 - \cos x + \operatorname{sen} 2x + 4 \operatorname{sen} 3x$$

es un polinomio trigonométrico de orden 3.

Centre su atención en el espacio vectorial $\mathscr{C}[-\pi, \pi]$ con el producto interno

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x) g(x) \, dx$$

Los polinomios trigonométricos de la forma en la ecuación (1) son combinaciones lineales del conjunto

$$\mathcal{B} = \{1, \cos x, \ldots, \cos nx, \operatorname{sen} x, \ldots, \operatorname{sen} nx\}$$

La mejor aproximación a una función f en $\mathscr{C}[-\pi, \pi]$ mediante un polinomio trigonométrico de orden n será por tanto $\operatorname{proy}_W(f)$, donde $W = \operatorname{gen}(\mathcal{B})$. Es evidente que $\mathcal{B}$ es un conjunto ortogonal y, por tanto, una base para W. La verificación de este hecho involucra demostrar que cualesquiera dos funciones distintas en $\mathcal{B}$ son ortogonales con respecto al producto interno dado. El ejemplo 7.43 presenta algunos de los cálculos necesarios; en los ejercicios 17-19 se le pide efectuar los restantes.

Ejemplo 7.43

Demuestre que sen jx es ortogonal a cos kx en $\mathscr{C}[-\pi, \pi]$ para $j, k \geq 1$.

Solución Al usar una identidad trigonométrica, calcule del modo siguiente: si $j \neq k$, entonces

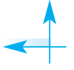
$$\begin{aligned} \int_{-\pi}^{\pi} \operatorname{sen} jx \cos kx \, dx &= \tfrac{1}{2}\int_{-\pi}^{\pi} [\operatorname{sen}(j+k)x + \operatorname{sen}(j-k)x] \, dx \\ &= -\tfrac{1}{2}\left[\frac{\cos(j+k)x}{j+k} + \frac{\cos(j-k)x}{j-k}\right]_{-\pi}^{\pi} \\ &= 0 \end{aligned}$$

pues la función coseno es periódica con periodo 2π.

Si $j = k$, entonces

$$\int_{-\pi}^{\pi} \operatorname{sen} kx \cos kx\, dx = \frac{1}{2k}[\operatorname{sen}^2 kx]_{-\pi}^{\pi} = 0$$

pues $\operatorname{sen} k\pi = 0$ para cualquier entero k.



Con la finalidad de encontrar la proyección ortogonal de una función f en $\mathscr{C}[-\pi, \pi]$ sobre el subespacio W generado por la base ortogonal $\mathcal{B}$, es necesario conocer los cuadrados de las normas de los vectores base. Por ejemplo, al usar una fórmula de ángulo medio, se tiene

$$\begin{aligned}\langle \operatorname{sen} kx, \operatorname{sen} kx\rangle &= \int_{-\pi}^{\pi} \operatorname{sen}^2 kx\, dx \\ &= \tfrac{1}{2}\int_{-\pi}^{\pi} (1 - \cos 2kx)\, dx \\ &= \tfrac{1}{2}\left[x - \frac{\operatorname{sen} 2kx}{2k}\right]_{-\pi}^{\pi} \\ &= \pi\end{aligned}$$

En el ejercicio 20 se le pide demostrar que $\langle \cos kx, \cos kx\rangle = \pi$ y $\langle 1, 1\rangle = 2\pi$.

Ahora se tiene

$$\operatorname{proy}_W(f) = a_0 + a_1 \cos x + \cdots + a_n \cos nx + b_1 \operatorname{sen} x + \cdots + b_n \operatorname{sen} nx \qquad (2)$$

donde

$$\begin{aligned} a_0 &= \frac{\langle 1, f\rangle}{\langle 1, 1\rangle} = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\, dx \\ a_k &= \frac{\langle \cos kx, f\rangle}{\langle \cos kx, \cos kx\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \cos kx\, dx \qquad (3) \\ b_k &= \frac{\langle \operatorname{sen} kx, f\rangle}{\langle \operatorname{sen} kx, \operatorname{sen} kx\rangle} = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \operatorname{sen} kx\, dx \end{aligned}$$

para $k \geq 1$. La aproximación a f dada por las ecuaciones (2) y (3) se llama ***aproximación de Fourier de n-ésimo orden*** a f sobre $[-\pi, \pi]$. Los coeficientes $a_0, a_1, \ldots, a_n, b_1, \ldots, b_n$ se llaman ***coeficientes de Fourier*** de f.

Ejemplo 7.44

Encuentre la aproximación de Fourier de cuarto orden a $f(x) = x$ sobre $[-\pi, \pi]$.

Solución Al usar las fórmulas (3), se obtiene

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} x\, dx = \frac{1}{2\pi}\left[\frac{x^2}{2}\right]_{-\pi}^{\pi} = 0$$

y para $k \geq 1$, la integración por partes produce

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x \cos kx\, dx = \frac{1}{\pi}\left[\frac{x}{k}\operatorname{sen} kx + \frac{1}{k^2}\cos kx\right]_{-\pi}^{\pi} = 0$$

Jean-Baptiste Joseph Fourier (1768-1830) fue un matemático y físico francés que ganó renombre por sus investigaciones acerca de la teoría del calor. En su histórica solución a la llamada ecuación térmica, introdujo técnicas relacionadas con lo que ahora se conoce como series de Fourier, una herramienta ampliamente usada en muchas ramas de las matemáticas, la física y la ingeniería. Fourier fue un activista político durante la Revolución francesa y se convirtió en favorito de Napoleón, a quien acompañó en su campaña a Egipto en 1798. Más tarde, Napoleón nombró a Fourier prefecto de Isère, donde supervisó muchos importantes proyectos de ingeniería. En 1808, Fourier se convirtió en barón. Una placa en la Torre Eiffel lo honra.

y

$$\begin{aligned} b_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} x \operatorname{sen} kx\,dx = \frac{1}{\pi}\left[-\frac{x}{k}\cos kx + \frac{1}{k^2}\operatorname{sen} kx\right]_{-\pi}^{\pi} \\ &= \frac{1}{\pi}\left[\frac{-\pi\cos k\pi - \pi\cos(-k\pi)}{k}\right] \\ &= \begin{cases} -\dfrac{2}{k} & \text{si } k \text{ es par} \\ \dfrac{2}{k} & \text{si } k \text{ es impar} \end{cases} \\ &= \frac{2(-1)^{k+1}}{k} \end{aligned}$$

Se tiene que la aproximación de Fourier de cuarto orden a $f(x) = x$ sobre $[-\pi, \pi]$ es

$$2(\operatorname{sen} x - \tfrac{1}{2}\operatorname{sen} 2x + \tfrac{1}{3}\operatorname{sen} 3x - \tfrac{1}{4}\operatorname{sen} 4x)$$

La figura 7.27 muestra las primeras cuatro aproximaciones de Fourier a $f(x) = x$ sobre $[-\pi, \pi]$.

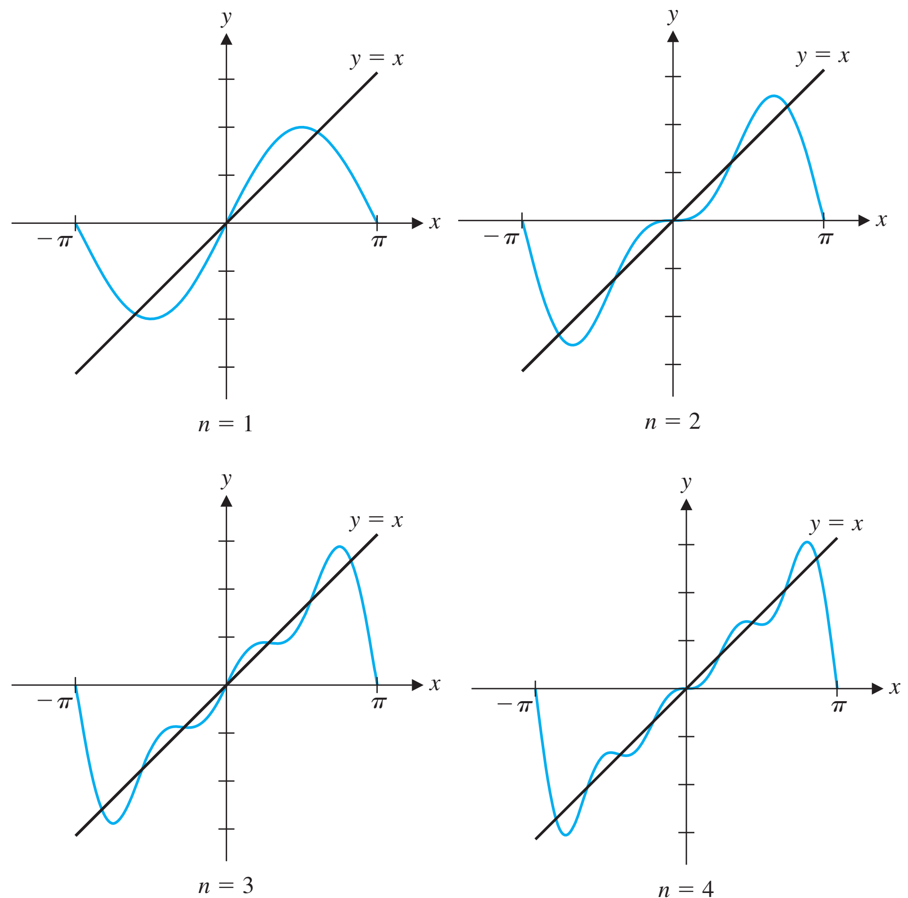
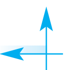


Figura 7.27



Claramente puede ver que las aproximaciones de la figura 7.27 mejoran, un hecho que puede confirmarse al calcular el error cuadrático medio en cada caso. Conforme aumenta el orden de las aproximaciones de Fourier, puede demostrarse que este error tiende a cero. Entonces el polinomio trigonométrico se convierte en una *serie infinita* y se escribe

$$f(x) = a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \operatorname{sen} kx)$$

Ésta se llama ***serie de Fourier*** de f sobre $[-\pi, \pi]$.

Códigos de corrección de error

Considere el código de triple repetición

$$C = \{\mathbf{c}_0, \mathbf{c}_1\} = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

Si ocurren uno o dos errores en la transmisión de alguno de dichos vectores código, el vector resultante no puede ser otro vector en C. De modo que C puede detectar hasta dos errores. Por ejemplo, si ocurren errores en la primera y segunda entradas cuando se transmite $\mathbf{c}_0$, entonces se recibe el vector

$$\mathbf{c}' = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Sin embargo, el receptor no tiene forma de corregir el error, pues $\mathbf{c}'$ también resultaría si ocurriera un solo error durante la transmisión de $\mathbf{c}_1$. Pero cualquier error *individual* puede corregirse, pues el vector resultante puede surgir sólo de una forma. Por ejemplo, si se recibe

$$\mathbf{c}'' = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

y se sabe que cuando mucho ocurrió un error, entonces el vector original debió ser $\mathbf{c}_0$, pues $\mathbf{c}''$ no puede surgir de $\mathbf{c}_1$ vía un solo error.

Ahora se generalizarán estas ideas. Como verá, la noción de distancia de Hamming tiene un papel crucial en la definición.

Definición Sea C un código (binario). La ***distancia mínima*** de C es la menor distancia de Hamming entre dos vectores distintos cualesquiera en C. Esto es,

$$d(C) = \text{mín}\{d_H(\mathbf{x}, \mathbf{y}) : \mathbf{x} \neq \mathbf{y} \text{ en } C\}$$

Claramente, la distancia mínima del anterior código de triple repetición C es 3.

Ejemplo 7.45

Encuentre la distancia mínima del código

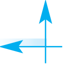
$$C = \{\mathbf{c}_0, \mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3\}$$

donde $$\mathbf{c}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{c}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{c}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{c}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Solución Es necesario calcular la distancia de Hamming entre cada par de distintos vectores. [Existen cuatro vectores, de modo que existen $\binom{4}{2} = 6$ pares.] Se encuentra que:

$$\mathrm{d}_H(\mathbf{c}_0, \mathbf{c}_1) = 2 \qquad \mathrm{d}_H(\mathbf{c}_0, \mathbf{c}_2) = 2 \qquad \mathrm{d}_H(\mathbf{c}_0, \mathbf{c}_3) = 4$$
$$\mathrm{d}_H(\mathbf{c}_1, \mathbf{c}_2) = 4 \qquad \mathrm{d}_H(\mathbf{c}_1, \mathbf{c}_3) = 2 \qquad \mathrm{d}_H(\mathbf{c}_2, \mathbf{c}_3) = 2$$

Por tanto, $\mathrm{d}(C) = 2$.



Es posible representar geométricamente las nociones de distancia mínima y de corrección de error. En el caso del código de triple repetición C, se tiene un subconjunto (en realidad, un subespacio) de $\mathbb{Z}_2^3$. Puede representar los vectores en $\mathbb{Z}_2^3$ como los vértices de un cubo unitario, como se muestra en la figura 7.28(a). La distancia de Hamming entre cualesquiera dos vectores $\mathbf{x}$ y $\mathbf{y}$ es justo el número de aristas en una trayectoria más corta de $\mathbf{x}$ a $\mathbf{y}$. El código C corresponde a dos de dichos vértices, $\mathbf{c}_0$ y $\mathbf{c}_1$. El hecho de que $\mathrm{d}(C) = 3$ corresponde al hecho de que $\mathbf{c}_0$ y $\mathbf{c}_1$ están separados tres unidades, como se muestra en la figura 7.28(b). Si un vector $\mathbf{x}$ recibido está dentro de una unidad de cualquiera de dichos vectores código y se sabe que cuando mucho ocurrió un error, puede decodificar correctamente $\mathbf{x}$ como el vector código *más cercano*. En la figura 7.28(b), $\mathbf{x}$ se decodificaría como $\mathbf{c}_0$ y $\mathbf{y}$ se decodificaría como $\mathbf{c}_1$. Esto concuerda con el hecho de que C puede corregir errores individuales mas no dobles.

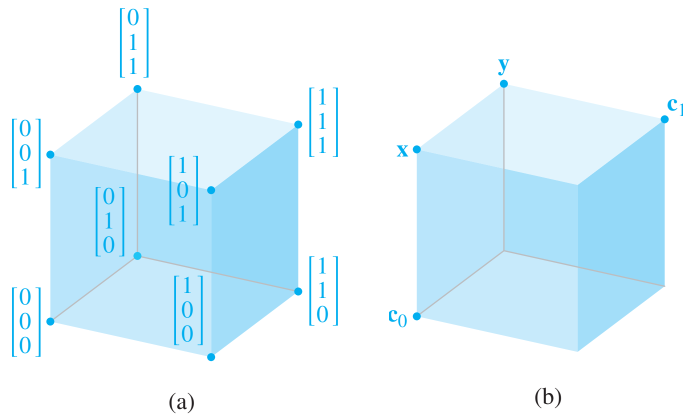


Figura 7.28

En el ejercicio 41 se le pide dibujar una imagen que ilustre la situación del ejemplo 7.45. En general, no es posible dibujar imágenes de $\mathbb{Z}_2^n$, pero es útil una analogía euclidiana. Si un código puede corregir hasta k errores, piense en los vectores código como en los centros de esferas de radio k. Los vectores código en ellos mismos están separados por al menos d unidades. Entonces, si un vector $\mathbf{x}$ recibido está dentro de una de dichas esferas, se decodificará como el vector correspondiente al centro de dicha esfera. En la figura 7.29, $\mathbf{x}$ se decodificará como $\mathbf{c}_0$. Este proceso se conoce como ***decodificación por el vecino más próximo***.

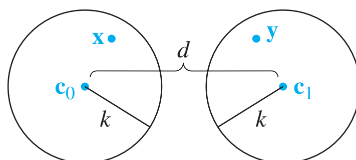


Figura 7.29

La figura 7.29 ilustra que si un código puede corregir k errores, entonces las "esferas" con centro en los vectores código no pueden tocarse o traslaparse; esto es: se debe tener $d > 2k$. Esto evidencia ser correcto, como se precisará ahora. Se dice que un código ***detecta k errores*** si, para cada vector de código $\mathbf{c}$ y cada vector $\mathbf{c}'$ que se obtiene al cambiar hasta k entradas de $\mathbf{c}$, $\mathbf{c}'$ no es un vector de código. Se dice que un código ***corrige k erro-***

res si, para cada vector de código $\mathbf{c}$ y cada vector $\mathbf{c}'$ que se obtiene al cambiar hasta k entradas de $\mathbf{c}$, la decodificación por el vecino más próximo de $\mathbf{c}'$ produce $\mathbf{c}$.

Teorema 7.20 Sea C un código (binario) con distancia mínima d.

a. C detecta k errores si y sólo si $d \geq k + 1$.
b. C corrige k errores si y sólo si $d \geq 2k + 1$.

Demostración (a) Suponga que $d \geq k + 1$ y sea $\mathbf{c}$ un vector en C. Si hasta k errores se introducen en $\mathbf{c}$, entonces el vector resultante $\mathbf{c}'$ tiene la propiedad de que $d_H(\mathbf{c}, \mathbf{c}') \leq k$. Pero entonces $\mathbf{c}'$ no puede ser un vector de código, pues si lo fuera, se tendría

$$k + 1 \leq d \leq d_H(\mathbf{c}, \mathbf{c}') \leq k$$

lo que es imposible.



Por el contrario, si C puede detectar hasta k errores, entonces la distancia mínima entre cualesquiera dos vectores de código debe ser mayor que k. (¿Por qué?) Se tiene que $d \geq k + 1$.

(b) Suponga que $d \geq 2k + 1$ y sea $\mathbf{c}$ un vector en C. Como en la demostración de la propiedad (a), sea $\mathbf{c}'$ un vector tal que $d_H(\mathbf{c}, \mathbf{c}') \leq k$. Sea $\mathbf{b}$ otro vector en C. Entonces $d_H(\mathbf{c}, \mathbf{b}) \geq d \geq 2k + 1$, de modo que, por la desigualdad del triángulo,

$$d_H(\mathbf{c}, \mathbf{c}') + d_H(\mathbf{c}', \mathbf{b}) \geq d_H(\mathbf{c}, \mathbf{b}) \geq 2k + 1$$

Por tanto,

$$d_H(\mathbf{c}', \mathbf{b}) \geq 2k + 1 - d_H(\mathbf{c}, \mathbf{c}') \geq 2k + 1 - k = k + 1 > d_H(\mathbf{c}', \mathbf{c})$$

De modo que $\mathbf{c}'$ está más cerca de $\mathbf{c}$ que de $\mathbf{b}$, y la decodificación por el vecino más próximo decodifica correctamente $\mathbf{c}'$ como $\mathbf{c}$.

Por el contrario, suponga que C puede corregir hasta k errores. Se demostrará que si $d < 2k + 1$ (esto es, $d \leq 2k$), entonces se obtiene una contradicción. Para hacer esto, se encontrará un vector código $\mathbf{c}$ y un vector $\mathbf{c}'$ tal que $d_H(\mathbf{c}, \mathbf{c}') \leq k$, pero la decodificación por el vecino más próximo decodifica $\mathbf{c}'$ como el vector de código *equivocado* $\mathbf{b} \neq \mathbf{c}$.

Sean $\mathbf{b}$ y $\mathbf{c}$ cualesquiera vectores de código en C tales que

$$d_H(\mathbf{b}, \mathbf{c}) = d \leq 2k$$

No hay daño al suponer que dichos d errores ocurren en las primeras d entradas de $\mathbf{b}$. (De otro modo, puede permutar las entradas de todos los vectores hasta que esto sea verdadero.) Si supone que los vectores de código en C tienen longitud n, construya un vector $\mathbf{c}'$ en $\mathbb{Z}_2^n$ del modo siguiente. Haga que $\mathbf{c}'$ concuerde con $\mathbf{b}$ en las primeras k entradas, concuerde con $\mathbf{c}$ en las siguientes $d - k$ entradas (¿por qué $d \geq k$?) y concuerde tanto con $\mathbf{b}$ como con $\mathbf{c}$ en las últimas $n - d$ entradas. En otras palabras, las entradas de $\mathbf{c}'$ satisfacen

$$c_i' = \begin{cases} b_i \neq c_i & \text{si } i = 1, \ldots, k \\ c_i \neq b_i & \text{si } i = k + 1, \ldots, d \\ b_i = c_i & \text{si } i = d + 1, \ldots, n \end{cases}$$



Ahora $d_H(\mathbf{c}, \mathbf{c}') = k$ y $d_H(\mathbf{c}', \mathbf{b}) = d - k \leq k$. (¿Por qué?) En consecuencia, $d_H(\mathbf{c}', \mathbf{b}) \leq d_H(\mathbf{c}', \mathbf{c})$, de modo que se tiene igualdad y es imposible decidir si $\mathbf{c}'$ debe decodificarse como $\mathbf{b}$ o $\mathbf{c}$, o la desigualdad es estricta y $\mathbf{c}'$ se decodificará incorrectamente como $\mathbf{b}$. En cualquier caso, se demostró que C no puede corregir k errores, lo que contradice la hipótesis. Se concluye que $d \geq 2k + 1$.

Algunos libros llaman a tal código un código $(n,2^k,d)$ o, de manera más general, un código (n, M, d), donde n es la longitud de los vectores, M es el número de vectores de código y d es la distancia mínima.

En el caso de un código lineal, se tiene la siguiente notación: si un código lineal (n, k) tiene distancia mínima d, se le conoce como un ***código (n, k, d)***. Por ejemplo, el código del ejemplo 7.45 es un código (4, 2, 2). Los códigos lineales tienen la ventaja de que su distancia mínima puede determinarse con facilidad. En el ejercicio 42 se le pide demostrar que la distancia mínima de un código lineal es la misma que el peso mínimo de un vector de código distinto de cero. También es posible determinar $\text{d}(C)$ al examinar una matriz de control de paridad para C.

Teorema 7.21 Sea C un código lineal (n, k) con matriz de verificación de paridad P. Entonces la distancia mínima de C es el entero más pequeño d para el cual P tiene d columnas linealmente dependientes.

Demostración Suponga que $\text{d}(C) = d$. La matriz de control de paridad P es una matriz de $(n - k) \times n$ con la propiedad de que, para cualquier vector $\mathbf{x}$ en $\mathbb{Z}_2^n$, $P\mathbf{x} = \mathbf{0}$ si y sólo si $\mathbf{x}$ está en C. Como se le pedirá demostrar en el ejercicio 42, C contiene un vector $\mathbf{c}$ de peso d. Entonces $P\mathbf{c}$ es una combinación lineal de exactamente d columnas de P. Pero, puesto que $P\mathbf{c} = \mathbf{0}$, esto implica que algún conjunto de d columnas de P es linealmente dependiente. Por otra parte, suponga que algún conjunto de $d - 1$ columnas de P es linealmente dependiente, por decir,

$$\mathbf{p}_{i_1} + \mathbf{p}_{i_2} + \cdots + \mathbf{p}_{i_{d-1}} = \mathbf{0}$$

Sea $\mathbf{x}$ un vector en $\mathbb{Z}_2^n$ con números 1 en las posiciones $i_1, \ldots, i_{d-1}$ y ceros en cualquier otra parte. Entonces $\mathbf{x}$ es un vector de peso $d - 1$ tal que $P\mathbf{x} = \mathbf{0}$. Por tanto, $\mathbf{x}$ es un vector de código de peso $d - 1 < d = \text{d}(C)$. Esto es imposible, por el ejercicio 42, así que se deduce que $\text{rank}(P) = d - 1$.

Por el contrario, suponga que cualesquiera $d - 1$ columnas de P son linealmente independientes pero algún conjunto de d columnas de P es linealmente dependiente. Puesto que $P\mathbf{x}$ es una combinación lineal de dichas columnas de P que corresponde a las posiciones de los 1 en $\mathbf{x}$, $P\mathbf{x} \neq \mathbf{0}$ para cualquier vector $\mathbf{x}$ de peso $d - 1$ o menor. Por tanto, no existen vectores código distintos de cero de peso menor que d. Pero algún conjunto de d columnas de P es linealmente dependiente, de modo que existe un vector $\mathbf{x}$ de peso d tal que $P\mathbf{x} = \mathbf{0}$. En consecuencia, esta $\mathbf{x}$ es un vector código de peso d. Por el ejercicio 42 nuevamente, se deduce que $\text{d}(C) = d$.

Ejemplo 7.46 Demuestre que todos los códigos de Hamming tienen distancia mínima 3.

Solución Recuerde que el código de Hamming (n, k) tiene una matriz de control de paridad de $(n - k) \times n$ cuyas columnas son todos los vectores distintos de cero de $\mathbb{Z}_2^{n-k}$, ordenados de modo que la matriz identidad ocupa las últimas $n - k$ columnas. Por ejemplo, el código de Hamming (7, 4) tiene matriz de control de paridad

$$P = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Siempre es posible encontrar tres columnas linealmente dependientes: sólo tome las columnas correspondientes a $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_1 + \mathbf{e}_2$. (En la matriz anterior, serían las columnas 5, 6 y 1, respectivamente.) Pero cualesquiera dos columnas son linealmente independientes. Por el Teorema 7.21, esto significa que los códigos de Hamming tienen distancia mínima 3.

El ejemplo 7.46, combinado con el Teorema 7.20, dice que los códigos de Hamming son todos de corrección de error individual. El otro tipo principal de código lineal que se consideró es la familia de códigos de Reed-Muller. Los mismos son capaces de corregir muchos errores, que es una de las razones por las que se eligen para transmitir fotografías desde el espacio.

Ejemplo 7.47

Demuestre que el código R_n de Reed-Muller tiene distancia mínima 2^{n-1} para $n \geq 1$.

Solución Por el Teorema 6.35, todo vector en R_n, excepto **0** y **1**, tiene peso 2^{n-1}. Dado que **1** tiene peso 2^n, esto significa que el peso mínimo de un vector de código distinto de cero en R_n es 2^{n-1}. Por tanto, $d(R_n) = 2^{n-1}$, por el ejercicio 42.

El *Mariner 9* usó el código R_5 de Reed-Muller, cuya distancia mínima es $2^4 = 16$. Por el Teorema 1, este código puede corregir k errores, donde $2k + 1 \leq 16$. El valor más grande de k para el cual esta desigualdad es verdadera es $k = 7$. Por tanto, R_5 no sólo contiene exactamente el número correcto de vectores de código para transmitir 64 tonos de gris, sino también es capaz de corregir hasta 7 errores, lo que lo hace muy confiable. ¡Esto explica por qué las imágenes transmitidas por el *Mariner 9* eran tan claras!

Ejercicios 7.5



Aproximación de funciones

En los ejercicios 1-4, encuentre la mejor aproximación lineal a f sobre el intervalo $[-1, 1]$.

1. $f(x) = x^2$ **2.** $f(x) = x^2 + 2x$

3. $f(x) = x^3$ **4.** $f(x) = \operatorname{sen}(\pi x/2)$

En los ejercicios 5 y 6, encuentre la mejor aproximación cuadrática a f sobre el intervalo $[-1, 1]$.

5. $f(x) = |x|$ **6.** $f(x) = \cos(\pi x/2)$

7. Aplique el proceso de Gram-Schmidt a la base $\{1, x\}$ para construir una base ortogonal para $\mathscr{P}_1[0, 1]$.

8. Aplique el proceso de Gram-Schmidt a la base $\{1, x, x^2\}$ para construir una base ortogonal para $\mathscr{P}_2[0, 1]$.

En los ejercicios 9-12, encuentre la mejor aproximación lineal a f sobre el intervalo $[0, 1]$.

9. $f(x) = x^2$ **10.** $f(x) = \sqrt{x}$

11. $f(x) = e^x$ **12.** $f(x) = \operatorname{sen}(\pi x/2)$

En los ejercicios 13-16, encuentre la mejor aproximación cuadrática a f sobre el intervalo $[0, 1]$.

13. $f(x) = x^3$ **14.** $f(x) = \sqrt{x}$

15. $f(x) = e^x$ **16.** $f(x) = \operatorname{sen}(\pi x/2)$

17. Demuestre que 1 es ortogonal a $\cos kx$ y $\operatorname{sen} kx$ en $\mathscr{C}[-\pi, \pi]$ para $k \geq 1$.

18. Demuestre que $\cos jx$ es ortogonal a $\cos kx$ en $\mathscr{C}[-\pi, \pi]$ para $j \neq k, j, k \geq 1$.

19. Demuestre que $\operatorname{sen} jx$ es ortogonal a $\operatorname{sen} kx$ en $\mathscr{C}[-\pi, \pi]$ para $j \neq k, j, k \geq 1$.

20. Demuestre que $\|1\|^2 = 2\pi$ y $\|\cos kx\|^2 = \pi$ en $\mathscr{C}[-\pi, \pi]$.

En los ejercicios 21 y 22, encuentre la aproximación de Fourier de tercer orden a f sobre $[-\pi, \pi]$.

21. $f(x) = |x|$ **22.** $f(x) = x^2$

En los ejercicios 23-26 encuentre los coeficientes de Fourier a_0, a_k *y* b_k *de f sobre* $[-\pi, \pi]$.

23. $f(x) = \begin{cases} 0 & \text{si } -\pi \leq x < 0 \\ 1 & \text{si } 0 \leq x \leq \pi \end{cases}$

24. $f(x) = \begin{cases} -1 & \text{si } -\pi \leq x < 0 \\ 1 & \text{si } 0 \leq x \leq \pi \end{cases}$

25. $f(x) = \pi - x$ **26.** $f(x) = |x|$

Recuerde que una función f es una ***función par*** *si* $f(-x) = f(x)$ *para toda x; f se llama* ***función impar*** *si* $f(-x) = -f(x)$ *para toda x.*

27. (a) Demuestre que $\displaystyle\int_{-\pi}^{\pi} f(x)\,dx = 0$ si f es una función impar.

(b) Demuestre que los coeficientes de Fourier a_k son todos cero si f es impar.

28. **(a)** Demuestre que $\int_{-\pi}^{\pi} f(x)\,dx = 2\int_{0}^{\pi} f(x)\,dx$ si f es una función par.

(b) Demuestre que los coeficientes de Fourier b_k son todos cero si f es par.

Códigos de corrección de error

Encuentre la distancia mínima de los códigos en los ejercicios 29-34.

29. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\}$

30. $C = \left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$

31. El código de paridad par E_n

32. El código de repetición de n veces, Rep_n

33. El código con matriz de control de paridad $P = [I \mid A]$, donde

$$A = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

34. El código con matriz de control de paridad

$$P = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

En los ejercicios 35 y 36, calcule la distancia mínima del código C y decodifique los vectores **u**, **v** *y* **w** *usando la decodificación por el vecino más próximo.*

35. $C = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}, \mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

36. C tiene matriz generadora

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \mathbf{u} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

En los ejercicios 37-40, construya un código lineal (n, k, d) o demuestre que no existe tal código.

37. $n = 8, k = 1, d = 8$

38. $n = 8, k = 2, d = 8$

39. $n = 8, k = 5, d = 5$

40. $n = 8, k = 4, d = 4$

41. Dibuje una imagen (similar a la figura 7.28) para ilustrar el ejemplo 7.45.

42. Sea C un código lineal. Demuestre que la distancia mínima de C es igual al peso mínimo de un vector de código distinto de cero.

43. Demuestre que $d - 1 \leq n - k$ para cualquier código lineal (n, k, d).

44. Sea C un código lineal (n, k, d) con matriz de control de paridad P. Pruebe que $d = n - k + 1$ si y sólo si todas las $n - k$ columnas de P son linealmente independientes.

Repaso del capítulo

Definiciones y conceptos clave

base ortogonal, 560
base ortonormal, 560
conjunto ortogonal de vectores, 560
conjunto ortonormal de vectores, 560
descomposición de valor singular (DVS), 616
desigualdad de Cauchy-Schwarz, 562
desigualdad del triángulo, 563
distancia, 558
distancia de Hamming, 577
esfera unitaria, 558
espacio con producto interno, 554
espacio lineal normado, 575
error de mínimos cuadrados, 595
matriz bien condicionada, 584
matriz mal condicionada, 584
norma, 575
norma de Frobenius, 579
norma de Hamming, 577
norma euclidiana (2-norma), 576
norma matricial, 579
norma máx (∞-norma, norma uniforme), 576
norma suma (1-norma), 575
número de condición de una matriz, 584
operador norma, 582
producto interno, 554
proyección ortogonal, 561, 606

Preguntas de repaso

1. Marque cada uno de los siguientes enunciados como verdadero o falso:

(a) Si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, entonces $\langle \mathbf{u}, \mathbf{v} \rangle = u_1 v_1 + \pi u_2 v_2$ define un producto interno sobre $\mathbb{R}^2$.

(b) Si $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ y $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, entonces $\langle \mathbf{u}, \mathbf{v} \rangle = 4u_1v_1 - 2u_1v_2 - 2u_2v_1 + 4u_2v_2$ define un producto interno sobre $\mathbb{R}^2$.

(c) $\langle A, B \rangle = \text{tr}(A) + \text{tr}(B)$ define un producto interno sobre M_{22}.

(d) Si $\mathbf{u}$ y $\mathbf{v}$ son vectores en un espacio con producto interno con $\|\mathbf{u}\| = 4$, $\|\mathbf{v}\| = \sqrt{5}$ y $\langle \mathbf{u}, \mathbf{v} \rangle = 2$, entonces $\|\mathbf{u} + \mathbf{v}\| = 5$.

(e) La norma suma, norma máx y la norma euclidiana sobre $\mathbb{R}^n$ son todas iguales a la función valor absoluto cuando $n = 1$.

(f) Si una matriz A está bien condicionada, entonces cond(A) es pequeña.

(g) Si cond(A) es pequeña, entonces la matriz A está bien condicionada.

(h) Todo sistema lineal tiene una solución única por mínimos cuadrados.

(i) Si A es una matriz con columnas ortonormales, entonces la matriz estándar de una proyección ortogonal sobre el espacio columna de A es $P = AA^T$.

(j) Si A es una matriz simétrica, entonces los valores singulares de A son iguales que los eigenvalores de A.

En las preguntas 2-4, determine si la definición proporciona un producto interno.

2. $\langle p(x), q(x) \rangle = p(0)q(1) + p(1)q(0)$ para $p(x), q(x)$ en $\mathscr{P}_1$

3. $\langle A, B \rangle = \text{tr}(A^T B)$ para A, B en M_{22}

4. $\langle f, g \rangle = \left(\max_{0 \le x \le 1} f(x)\right)\left(\max_{0 \le x \le 1} g(x)\right)$ para f, g en $\mathscr{C}[0, 1]$

En las preguntas 5 y 6, calcule la cantidad indicada usando el producto interno especificado.

5. $\|1 + x + x^2\|$ si $\langle a_0 + a_1x + a_2x^2, b_0 + b_1x + b_2x^2 \rangle = a_0b_0 + a_1b_1 + a_2b_2$

6. $d(x, x^2)$ si $\langle p(x), q(x) \rangle = \int_0^1 p(x)q(x)\,dx$

En las preguntas 7 y 8, construya un conjunto ortogonal de vectores al aplicar el proceso de Gram-Schmidt al conjunto dado de vectores, usando el producto interno especificado.

7. $\left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$ si $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^T A \mathbf{v}$, donde $A = \begin{bmatrix} 6 & 4 \\ 4 & 6 \end{bmatrix}$

8. $\{1, x, x^2\}$ si $\langle p(x), q(x) \rangle = \int_0^1 p(x)q(x)\,dx$

En las preguntas 9 y 10, determine si la definición proporciona una norma.

9. $\|\mathbf{v}\| = \mathbf{v}^T\mathbf{v}$ para $\mathbf{v}$ en $\mathbb{R}^n$

10. $\|p(x)\| = |p(0)| + |p(1) - p(0)|$ para $p(x)$ en $\mathscr{P}_1$

11. Demuestre que la matriz $A = \begin{bmatrix} 1 & 0.1 & 0.11 \\ 0.1 & 0.11 & 0.111 \\ 0.11 & 0.111 & 0.1111 \end{bmatrix}$ está mal condicionada.

12. Pruebe que si Q es una matriz ortogonal de $n \times n$, entonces su norma de Frobenius es $\|Q\|_F = \sqrt{n}$.

13. Encuentre la recta de mejor ajuste a través de los puntos $(1, 2)$, $(2, 3)$, $(3, 5)$ y $(4, 7)$.

14. Encuentre la solución por mínimos cuadrados de

$$\begin{bmatrix} 1 & 2 \\ 1 & 0 \\ 2 & -1 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 3 \end{bmatrix}.$$

15. Encuentre la proyección ortogonal de $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ sobre el espacio columna de $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

16. Si $\mathbf{u}$ y $\mathbf{v}$ son vectores ortonormales, demuestre que $P = \mathbf{u}\mathbf{u}^T + \mathbf{v}\mathbf{v}^T$ es la matriz estándar de una proyección ortogonal sobre gen$(\mathbf{u}, \mathbf{v})$. [*Sugerencia:* demuestre que $P = A(A^TA)^{-1}A^T$ para alguna matriz A.]

En las preguntas 17 y 18, encuentre (a) los valores singulares, (b) una descomposición de valor singular y (c) la pseudoinversa de la matriz A.

17. $A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & -1 \end{bmatrix}$

18. $A = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \end{bmatrix}$

19. Si P y Q son matrices ortogonales para las cuales está definida PAQ, pruebe que PAQ tiene los mismos valores singulares que A.

20. Si A es una matriz cuadrada para la cual $A^2 = O$, pruebe que $(A^+)^2 = O$.

Apéndice A*

Notación matemática y métodos de demostración

Por favor, señor, quiero un poco más.
—Oliver
Charles Dickens, *Oliver Twist*

Quienquiera que entienda la notación algebraica lee de un vistazo en una ecuación resultados que se alcanzan aritméticamente sólo con gran trabajo y dolor.
—Augustin Cournot
Researches into the Mathematical Principles of the Theory of Wealth
Traducido por Nathaniel T. Bacon
Macmillan, 1897, p. 4

En este libro se hizo un esfuerzo para usar "idioma matemático" tanto como fuera posible y mantener la notación matemática al mínimo. Sin embargo, la notación matemática es una taquigrafía conveniente que puede simplificar enormemente la cantidad de escritura que debe realizarse. Más aún, se usa comúnmente en toda rama de las matemáticas, de modo que la habilidad de leer y escribir notación matemática es un ingrediente esencial de la comprensión matemática. Finalmente, existen algunos teoremas cuyas demostraciones se vuelven "obvias" si se usa la notación correcta.

Demostrar teoremas en matemáticas es tanto un arte como una ciencia. Para el principiante, con frecuencia es difícil saber cuál método usar para demostrar un teorema; existen muchos métodos, y cualquiera puede evidenciar ser el mejor. Para volverse hábil en las demostraciones, es importante estudiar tantos ejemplos como sea posible y tener mucha práctica.

Este apéndice resume la notación matemática básica que se aplica a los conjuntos. También se estudia la notación suma, una útil taquigrafía para lidiar con las sumas. Finalmente, se ilustran algunos métodos de demostración con ejemplos genéricos.

Notación de conjuntos

Un ***conjunto*** es una colección de objetos, llamados ***elementos*** (o ***miembros***) del conjunto. Los ejemplos de conjuntos incluyen el conjunto de todas las palabras en este texto, el conjunto de todos los libros en la biblioteca de su escuela, el conjunto de los enteros positivos y el conjunto de todos los vectores en el plano cuya ecuación es $2x + 3y - z = 0$.

Con frecuencia es posible mencionar los elementos de un conjunto, en cuyo caso es convencional encerrar la lista dentro de llaves. Por ejemplo, se tiene

$$\{1, 2, 3\}, \quad \{a, t, x, z\}, \quad \{2, 4, 6, \ldots, 100\}, \quad \left\{\frac{\pi}{4}, \frac{2\pi}{5}, \frac{\pi}{2}, \frac{4\pi}{7}, \ldots, \frac{5\pi}{6}\right\}$$



Note que los puntos suspensivos (. . .) denotan elementos omitidos cuando se presenta un patrón. (¿Cuál es el patrón en los últimos dos ejemplos?) Los conjuntos infinitos con frecuencia se expresan usando puntos suspensivos. Por ejemplo, el conjunto de los enteros positivos por lo general se denota mediante $\mathbb{N}$ o $\mathbb{Z}^+$, de modo que

$$\mathbb{N} = \mathbb{Z}^+ = \{1, 2, 3, \ldots\}$$

El conjunto de todos los enteros se denota mediante $\mathbb{Z}$, de modo que

$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$$

Dos conjuntos se consideran ***iguales*** si contienen exactamente los mismos elementos. El *orden* en el que se mencionan los elementos no importa, y las repeticiones no se cuentan. Por tanto,

*En el sitio Web para el estudiante puede encontrar ejercicios y respuestas seleccionadas con número impar para este apéndice.

$$\{1, 2, 3\} = \{2, 1, 3\} = \{1, 3, 2, 1\}$$

El símbolo $\in$ significa "es elemento de" o "está en", y el símbolo $\notin$ denota la negación; esto es, "no es elemento de" o "no está en". Por ejemplo,

$$5 \in \mathbb{Z}^+ \quad \text{pero} \quad 0 \notin \mathbb{Z}^+$$

Con frecuencia es más conveniente describir un conjunto en términos de una regla que satisfagan todos sus elementos. En tales casos es apropiada la ***notación de construcción de conjuntos***. El formato es

$$\{x : x \text{ satisface a } P\}$$

donde P representa una propiedad o una colección de propiedades que debe satisfacer el elemento x. Los dos puntos se pronuncian "tal que". Por ejemplo,

$$\{n : n \in \mathbb{Z}, n > 0\}$$

se lee como "el conjunto de todas las n tales que n es un entero y n es mayor que cero". Esta sólo es otra forma de describir los enteros positivos $\mathbb{Z}^+$. (También podría escribirse $\mathbb{Z}^+ = \{n \in \mathbb{Z} : n > 0\}$.)

El ***conjunto vacío*** es el conjunto sin elementos. Se denota $\varnothing$ o $\{\ \}$.

Ejemplo A.1

Describa con palabras los siguientes conjuntos:

(a) $A = \{n : n = 2k, k \in \mathbb{Z}\}$ (b) $B = \{m/n : m, n \in \mathbb{Z}, n \neq 0\}$
(c) $C = \{x \in \mathbb{R} : 4x^2 - 4x - 3 = 0\}$ (d) $D = \{x \in \mathbb{Z} : 4x^2 - 4x - 3 = 0\}$

Solución (a) A es el conjunto de números n que son múltiplos enteros de 2. Por tanto, A es el conjunto de todos los enteros pares.

(b) B es el conjunto de todas las expresiones de la forma m/n, donde m y n son enteros y n es distinto de cero. Este es el conjunto de *números racionales*, que usualmente se denota $\mathbb{Q}$. (Note que esta forma de describir $\mathbb{Q}$ produce muchas repeticiones; sin embargo, la convención, que se puntualizó anteriormente, es que sólo se incluya una vez cada elemento. Por tanto, esta expresión describe precisamente el conjunto de todos los números racionales.)

(c) C es el conjunto de todas las soluciones reales de la ecuación $4x^2 - 4x - 3 = 0$. Al factorizar o usar la fórmula cuadrática, se encuentra que las raíces de esta ecuación son $-\frac{1}{2}$ y $\frac{3}{2}$. (Verifique esto.) Por tanto,

$$C = \{-\tfrac{1}{2}, \tfrac{3}{2}\}$$

(d) A partir de la solución a (c) se ve que *no* existen soluciones para $4x^2 - 4x - 3 = 0$ en $\mathbb{R}$ que sean enteras. Por tanto, D es el conjunto vacío, que puede expresarse al escribir $D = \varnothing$.

John Venn (1834-1923) fue un matemático inglés que estudió en la Universidad de Cambridge y más tarde dio clases ahí. Trabajó principalmente en lógica matemática y es mejor conocido por inventar los diagramas de Venn.

Si todo elemento de un conjunto A también es un elemento de un conjunto B, entonces A es un ***subconjunto*** de B, lo que se denota $A \subseteq B$. Esta situación puede representarse esquemáticamente con el uso de un ***diagrama de Venn***, como se muestra en la figura A.1. (El rectángulo representa el *conjunto universal*, un conjunto suficientemente grande para contener todos los otros conjuntos en cuestión, en este caso, A y B.)

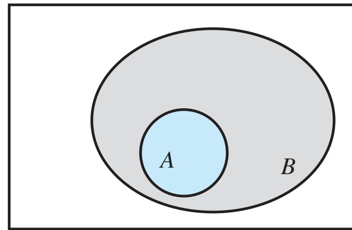


Figura A.1
$A \subseteq B$

Ejemplo A.2

(a) $\{1, 2, 3\} \subseteq \{1, 2, 3, 4, 5\}$

(b) $\mathbb{Z}^+ \subseteq \mathbb{Z} \subseteq \mathbb{R}$

(c) Sea A el conjunto de todos los enteros positivos cuyos dos últimos dígitos son 24 y sea B el conjunto de todos los enteros positivos que son divisibles entre 4. Entonces, si n está en A, es de la forma

$$n = 100k + 24$$

para algún entero k. (Por ejemplo, $36{,}524 = 100 \cdot 365 + 24$.) Pero entonces

$$n = 100k + 24 = 4(25k + 6)$$

de modo que $n/4 = 25k + 6$, que es un entero. En consecuencia, n es divisible entre 4, como también lo es B. Por tanto, $A \subseteq B$.

Es posible demostrar que dos conjuntos A y B son iguales al probar que cada uno es un subconjunto del otro. Esta estrategia es particularmente útil si los conjuntos se definen de manera abstracta o si no es fácil mencionar y comparar sus elementos.

Ejemplo A.3

Sea A el conjunto de todos los enteros positivos cuyos dos últimos dígitos forman un número que es divisible entre 4. En el caso de un número de un dígito, se toma su dígito de decenas como 0. Sea B el conjunto de todos los enteros positivos que son divisibles entre 4. Demuestre que $A = B$.

Solución Como en el ejemplo A.2(c), es fácil ver que $A \subseteq B$. Si n está en A, entonces es posible dividir el número m formado por sus dos últimos dígitos al escribir

$$n = 100k + m$$

para algún entero k. Pero, puesto que m es divisible entre 4, se tiene $m = 4r$ para algún entero r.

$$n = 100k + m = 100k + 4r = 4(25k + r)$$

de modo que n también es divisible entre 4. En consecuencia, $A \subseteq B$.

Para demostrar que $B \subseteq A$, sea n que está en B. Esto es, n es divisible entre 4. Suponga que $n = 4s$, donde s es un entero. Si m es el número formado por los dos últimos dígitos de n, entonces, como antes, $n = 100k + m$ para algún entero k. Pero ahora

$$m = n - 100k = 4s - 100k = 4(s - 25k)$$

lo que implica que m es divisible entre 4, pues $s - 25k$ es un entero. Por tanto, n está en A, y se demostró que $B \subseteq A$.

Puesto que $A \subseteq B$ y $B \subseteq A$, se debe tener $A = B$.

La ***intersección*** de los conjuntos A y B se denota $A \cap B$ y consiste de los elementos que A y B tienen en común. Esto es:

$$A \cap B = \{x : x \in A \quad \text{y} \quad x \in B\}$$

La figura A.2 muestra un diagrama de Venn para este caso. La ***unión*** de A y B se denota $A \cup B$ y consiste de los elementos que están en A o en B (o en ambos). Esto es,

$$A \cup B = \{x : x \in A \quad \text{o} \quad x \in B\}$$

Vea la figura A.3.

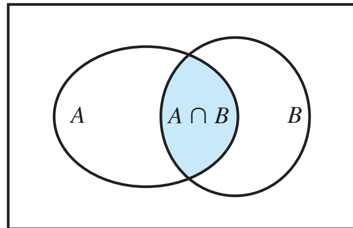


Figura A.2
$A \cap B$

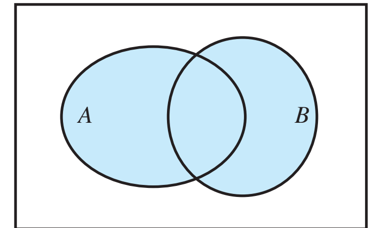


Figura A.3
$A \cup B$

Ejemplo A.4

Sea $A = \{n^2 : n \in \mathbb{Z}^+, 1 \leq n \leq 4\}$ y sea $B = \{n \in \mathbb{Z}^+ : n \leq 10 \text{ y } n \text{ es impar}\}$. Encuentre $A \cap B$ y $A \cup B$.

Solución Se ve que

$$A = \{1^2, 2^2, 3^2, 4^2\} = \{1, 4, 9, 16\} \qquad \text{y} \qquad B = \{1, 3, 5, 7, 9\}$$

Por tanto, $A \cap B = \{1, 9\}$ y $A \cup B = \{1, 3, 4, 5, 7, 9, 16\}$.

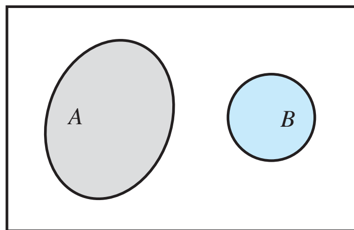


Figura A.4
Conjuntos disjuntos

Si $A \cap B = \varnothing$, entonces A y B se llaman ***conjuntos disjuntos***. (Vea la figura A.4.) Por ejemplo, el conjunto de enteros pares y el conjunto de enteros impares son disjuntos.

Notación suma

La notación suma es una taquigrafía conveniente para escribir una suma como

$$1 + 2 + 3 + \cdots + 100$$

donde se quiere eliminar casi todos los elementos. Como en la notación de conjuntos, los puntos suspensivos (. . .) significan que se estableció un patrón y que simplemente se quitaron algunos términos intermedios. En el ejemplo anterior se espera que los lectores reconozcan que se suman todos los enteros positivos de 1 a 100. Sin embargo, los puntos suspensivos pueden ser ambiguos. Por ejemplo, ¿cómo sería la suma

$$1 + 2 + \cdots + 64?$$

¿Es la suma de todos los enteros positivos desde 1 hasta 64 o sólo de las potencias de dos, $1 + 2 + 4 + 8 + 16 + 32 + 64$? Con frecuencia es más claro (y más corto) usar ***notación suma*** (o ***notación sigma***).

Σ es la letra griega mayúscula *sigma*, que corresponde a S (por "suma"). La notación suma la introdujo Fourier en 1820 y la comunidad matemática la adoptó rápidamente.

Es posible abreviar una suma de la forma

$$a_1 + a_2 + \cdots + a_n \tag{1}$$

como

$$\sum_{k=1}^{n} a_k \tag{2}$$

que dice que hay que sumar los términos a_k sobre todos los enteros k que van desde 1 hasta n. Una versión alternativa de esta expresión es

$$\sum_{1 \le k \le n} a_k$$

El subíndice k se llama ***índice de suma***. Es una "variable falsa" en el sentido de que no aparece en la suma real de la expresión (1). Por tanto, puede usar cualquier letra que quiera como el índice de suma (en tanto no aparezca ya en alguna otra parte de las expresiones que se están sumando). Por ende, la expresión (2) también puede escribirse como

$$\sum_{i=1}^{n} a_i$$

El índice de suma no necesita comenzar en 1. La suma $a_3 + a_4 + \cdots + a_{99}$ se convierte en

$$\sum_{k=3}^{99} a_k$$

aunque puede disponer que el índice comience en 1 al reescribir la expresión como

$$\sum_{k=1}^{97} a_{k+2}.$$

La clave para usar la notación suma de manera efectiva es poder reconocer los patrones.

Ejemplo A.5

Escriba las siguientes sumas usando la notación suma.

(a) $1 + 2 + 4 + \cdots + 64$ (b) $1 + 3 + 5 + \cdots + 99$ (c) $3 + 8 + 15 + \cdots + 99$

Solución (a) Esta expresión se reconoce como una suma de potencias de 2:

$$1 + 2 + 4 + \cdots + 64 = 2^0 + 2^1 + 2^2 + \cdots + 2^6$$

Por tanto, el índice de suma aparece como el exponente y se tiene $\sum_{k=0}^{6} 2^k$.

(b) Esta expresión es la suma de todos los enteros impares desde 1 hasta 99. Todo entero impar es de la forma $2k + 1$, de modo que la suma es

$$\begin{aligned} &1 + 3 + 5 + \cdots + 99 \\ &\quad = (2 \cdot 0 + 1) + (2 \cdot 1 + 1) + (2 \cdot 2 + 1) + \cdots + (2 \cdot 49 + 1) \\ &\quad = \sum_{k=0}^{49} (2k + 1) \end{aligned}$$

(c) Aquí el patrón es menos claro, pero un poco de reflexión revela que cada término es 1 menos que un cuadrado perfecto:

$$\begin{aligned} &3 + 8 + 15 + \cdots + 99 \\ &\quad = (2^2 - 1) + (3^2 - 1) + (4^2 - 1) + \cdots + (10^2 - 1) \\ &\quad = \sum_{k=2}^{10} (k^2 - 1) \end{aligned}$$

Ejemplo A.6

Reescriba cada una de las sumas en el ejemplo A.5 de modo que el índice suma comience en 1.

Solución (a) Si usa el cambio de variable $i = k + 1$, entonces, conforme k avanza de 0 a 6, i va de 1 a 7. Puesto que $k = i - 1$, se obtiene

$$\sum_{k=0}^{6} 2^k = \sum_{i=1}^{7} 2^{i-1}$$

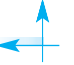
(b) Al usar la misma sustitución que en el inciso (a), se obtiene

$$\sum_{k=0}^{49} (2k + 1) = \sum_{i=1}^{50} (2(i - 1) + 1) = \sum_{i=1}^{50} (2i - 1)$$



(c) La sustitución $i = k - 2$ funcionará (inténtelo), pero es más fácil sólo agregar un término correspondiente a $k = 1$, pues $1^2 - 1 = 0$. Por tanto,

$$\sum_{k=2}^{10} (k^2 - 1) = \sum_{k=1}^{10} (k^2 - 1)$$



Cuando hay más de un índice de suma surgen múltiples sumas, como los hay en una matriz. La notación

$$\sum_{i,j=1}^{n} a_{ij} \tag{3}$$

significa sumar los términos a_{ij} conforme i y j cada uno varían independientemente desde 1 hasta n. La suma en la expresión (3) es equivalente a

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}$$

donde la suma es primero sobre j y luego sobre i (siempre se trabaja de dentro hacia afuera) o

$$\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}$$

donde se invierte el orden de la suma.

Ejemplo A.7

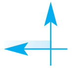
Escriba $\sum_{i,j=1}^{3} i^j$ usando ambos órdenes posibles de suma.

Solución

$$\begin{aligned}\sum_{i=1}^{3} \sum_{j=1}^{3} i^j &= \sum_{i=1}^{n} (i^1 + i^2 + i^3)\\ &= (1^1 + 1^2 + 1^3) + (2^1 + 2^2 + 2^3) + (3^1 + 3^2 + 3^3)\\ &= (1 + 1 + 1) + (2 + 4 + 8) + (3 + 9 + 27) = 56\end{aligned}$$

y

$$\sum_{j=1}^{3}\sum_{i=1}^{3} i^j = \sum_{j=1}^{n}(1^j + 2^j + 3^j)$$
$$= (1^1 + 2^1 + 3^1) + (1^2 + 2^2 + 3^2) + (1^3 + 2^3 + 3^3)$$
$$= (1 + 2 + 3) + (1 + 4 + 9) + (1 + 8 + 27) = 56$$



Comentario Desde luego, el valor de la suma en el ejemplo A. 7 es el mismo sin importar cuál orden de suma se elija, porque la suma es *finita*. También es posible considerar *sumas infinitas* (conocidas en cálculo como *series infinitas*), pero tales sumas no siempre tienen un valor y debe tener mucho cuidado cuando se reordenan o manipulan sus términos. Por ejemplo, suponga que se tiene

$$S = \sum_{k=0}^{\infty} 2^k$$

Entonces

$$S = 1 + 2 + 4 + 8 + \cdots$$
$$= 1 + 2(1 + 2 + 4 + \cdots)$$
$$= 1 + 2S$$

de donde se sigue que $S = -1$. Claramente esto no tiene sentido, ¡pues S es una suma de términos *no negativos*! (¿Dónde está el error?)

How to Solve It es el título de un libro del matemático George Pólya (1887-1985). Desde su publicación en 1945, *How to Solve It* ha vendido más de un millón de copias y se ha traducido a 17 idiomas. Pólya nació en Hungría, pero, debido a la situación política en Europa, se mudó a Estados Unidos en 1940. Posteriormente impartió cátedra en las Universidades Brown y Stanford, donde realizó investigación matemática y desarrolló una reputación bien merecida como destacado profesor. El Premio Pólya lo otorga anualmente la Society for Industrial and Applied Mathematics para grandes aportaciones en las áreas de las matemáticas cercanas a aquellas en las que trabajó Pólya. La Mathematical Association of America otorga anualmente la Cátedra Pólya a los matemáticos que demuestran la exposición de alta calidad por la cual se conoció a Pólya.

Métodos de demostración

La noción de demostración está en el corazón mismo de la matemática. Una cosa es saber *qué* es verdadero; otra muy distinta es saber *por qué* es verdadero y poder demostrar su veracidad mediante una secuencia de enunciados lógicamente conectados. La intención aquí no es tratar de enseñar cómo hacer demostraciones; usted se volverá cada vez mejor para hacer demostraciones al estudiar los ejemplos y practicando, algo que debe hacer con frecuencia mientras trabaja a lo largo de este libro. La intención de esta breve sección es simplemente ofrecer algunos ejemplos elementales de ciertos tipos de demostraciones. Las demostraciones de los teoremas en el texto proporcionarán más ilustraciones de "cómo resolverlo".

Hablando en general, las demostraciones matemáticas caen en dos categorías: ***demostraciones directas*** y ***demostraciones indirectas***. Muchos teoremas tienen la estructura "si P, entonces Q", donde P (la *hipótesis* o *premisa*) y Q (la *conclusión*) son enunciados que son o verdaderos o falsos. Tal implicación se denota $P \Rightarrow Q$. Una demostración directa procede al establecer una cadena de implicaciones

$$P \Rightarrow P_1 \Rightarrow P_2 \Rightarrow \cdots \Rightarrow P_n \Rightarrow Q$$

lo que conduce de P a Q.

Ejemplo A.8

Demuestre que cualesquiera dos cuadrados perfectos consecutivos difieren por un número impar. Esta instrucción puede replantearse como: "demuestre que si a y b son cuadrados perfectos consecutivos, entonces $a - b$ es impar". Por tanto, tiene la forma $P \Rightarrow Q$, con P como "a y b son cuadrados perfectos consecutivos" y Q como "$a - b$ es impar".

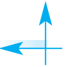
Solución Suponga que a y b son cuadrados perfectos consecutivos, con $a > b$. Entonces

$$a = (n + 1)^2 \quad \text{y} \quad b = n^2$$

para algún entero n. Pero ahora

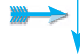
$$\begin{aligned} a - b &= (n + 1)^2 - n^2 \\ &= n^2 + 2n + 1 - n^2 \\ &= 2n + 1 \end{aligned}$$

de modo que $a - b$ es impar.



Existen dos tipos de demostraciones indirectas que pueden usarse para establecer un enunciado condicional de la forma $P \Rightarrow Q$. Una ***demostración por contradicción*** supone que la hipótesis P es verdadera, como en una demostración directa, pero luego supone que la conclusión Q es *falsa*. Entonces la estrategia es demostrar que esto no es posible (es decir: descartar la posibilidad de que la conclusión sea falsa) al encontrar una contradicción a la veracidad de P. Entonces se concluye que Q debe ser verdadera.

Ejemplo A.9



Sea n un entero positivo. Demuestre que si n^2 es par, también lo es n. (Tome algunos minutos para tratar de encontrar una demostración directa de esta afirmación; le ayudará a apreciar la demostración indirecta que sigue.)

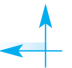
Solución Suponga que n es un entero positivo tal que n^2 es par. Ahora suponga que n no es par. Entonces n es impar, de modo que

$$n = 2k + 1$$

para algún entero k. Pero si es así, se tiene

$$n^2 = (2k + 1)^2 = 4k^2 + 4k + 1$$

de modo que n^2 es impar, pues es 1 más que el número par $4k^2 + 4k$. Esto contradice la hipótesis de que n^2 es par. Se concluye que la suposición de que n *no* era par debe ser falsa; en otras palabras, n debe ser par.



Estrechamente relacionada con el método de demostración por contradicción está la ***demostración por contrapositivo***. El *negativo* de un enunciado P es el enunciado "no es el caso que P", que se abrevia simbólicamente como $\neg P$ y se pronuncia "no P". Por ejemplo, si P es "n es par", entonces $\neg P$ es "no es el caso que n sea par"; en otras palabras "n es impar".

El *contrapositivo* del enunciado $P \Rightarrow Q$ es el enunciado $\neg Q \Rightarrow \neg P$. Un enunciado condicional $P \Rightarrow Q$ y su contrapositivo $\neg Q \Rightarrow \neg P$ son lógicamente equivalentes en el sentido de que son o ambos verdaderos o ambos son falsos. (Por ejemplo, si $P \Rightarrow Q$ es un teorema, entonces también lo es $\neg Q \Rightarrow \neg P$. Para ver esto, note que si la hipótesis $\neg Q$ es verdadera, entonces Q es falsa. La conclusión $\neg P$ no puede ser falsa porque, si lo fuera, entonces P sería verdadera y el teorema conocido $P \Rightarrow Q$ implicaría la veracidad de Q, lo que produciría una contradicción. Se concluye que $\neg P$ es verdadero y se demostró $\neg Q \Rightarrow \neg P$.) He aquí una demostración contrapositiva de la afirmación del ejemplo A.9.



Ejemplo A.10

Sea n un entero positivo. Demuestre que si n^2 es par, también lo es n.

Solución El contrapositivo del enunciado dado es

"Si n no es par, entonces n^2 no es par" o "si n es impar, también lo es n^2"

Para demostrar este contrapositivo, suponga que n es impar. Entonces $n = 2k + 1$ para algún entero k. Como antes, esto significa que $n^2 = (2k + 1)^2 = 4k^2 + 4k + 1$ es impar, lo que completa la demostración del contrapositivo. Puesto que el contrapositivo es verdadero, también lo es el enunciado original.

Aunque no se requiere un nuevo método de demostración para manejarlo, se considerará brevemente cómo demostrar un teorema "si y sólo si". Un enunciado de la forma "P si y sólo si Q" señala una *doble implicación*, que se denota $P \Leftrightarrow Q$. Para demostrar tal enunciado, debe demostrar $P \Rightarrow Q$ y $Q \Rightarrow P$. Para hacerlo, puede usar las técnicas descritas anteriormente donde sea adecuado. Es importante notar que la parte "*si*" de $P \Leftrightarrow Q$ es "P es Q", que es $Q \Rightarrow P$; la parte "sólo si" de $P \Leftrightarrow Q$ es "P sólo si Q", lo que significa $P \Rightarrow Q$. La implicación $P \Rightarrow Q$ en ocasiones se lee como "P es suficiente para Q" o "Q es necesaria para P"; $Q \Rightarrow P$ se lee "Q es suficiente para P" o "P es necesaria para Q". En conjunto, son $P \Leftrightarrow Q$, o "P es necesaria y suficiente para Q" y viceversa.

Ejemplo A.11

Un peón se coloca en un tablero de ajedrez y se le permite moverse un cuadrado a la vez, ya sea horizontal o verticalmente. El *recorrido de un peón* en un tablero de ajedrez es la trayectoria que sigue un peón, al moverse como se describe, que visita un cuadrado exactamente una vez, y comienza y termina en el mismo cuadrado. Demuestre que existe un recorrido de peón de un tablero de ajedrez de $n \times n$ si y sólo si n es par.

Solución [$\Leftarrow$] ("si") Suponga que n es par. Es fácil ver que la estrategia que se ilustra en la figura A.5 para un tablero de 6×6 siempre producirá un recorrido de peón.

[$\Rightarrow$] ("sólo si") Suponga que existe un recorrido de peón en un tablero de $n \times n$. Se dará una demostración por contradicción de que n debe ser par. Para este fin, suponga que n es impar. En cada movimiento, el peón se mueve a un cuadrado de diferente color. El número total de movimientos en su recorrido es n^2, que también es un número impar, de acuerdo con la demostración del ejemplo A.10. Por tanto, el peón debe terminar en un cuadrado del color opuesto al del cuadrado desde el cual comenzó. (¿Por qué?) Esto es imposible, pues el peón termina donde comenzó, así que hay una contradicción. Se concluye que n no puede ser impar; por tanto, n es par y la demostración está completa.

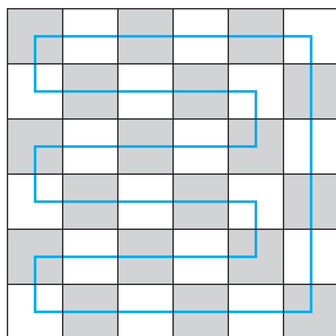


Figura A.5

Algunos teoremas afirman que varios enunciados son *equivalentes*. Esto significa que cada uno es verdadero si y sólo si todos los demás son verdaderos. Demostrar que n enunciados son equivalentes requiere $\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n^2 - n}{2}$ demostraciones "si y sólo si". Sin embargo, en la práctica, con frecuencia es más fácil establecer un "anillo" de n implicaciones que vincular todos los enunciados. La demostración del teorema fundamental de las matrices invertibles brinda un excelente ejemplo de este enfoque.

Apéndice B*

Inducción matemática

Las pulgas grandes tienen pulgas pequeñas en su espaldas que las muerden y las pulgas pequeñas tienen pulgas más pequeñas, y así ad infinitum

–Augustus De Morgan
A Budget of Paradoxes
Longman, Green, and Company, 1872, p. 377

La habilidad para reconocer patrones es una de las claves para triunfar en la resolución de problemas matemáticos. Considere el siguiente patrón:

$$1 = 1$$
$$1 + 3 = 4$$
$$1 + 3 + 5 = 9$$
$$1 + 3 + 5 + 7 = 16$$
$$1 + 3 + 5 + 7 + 9 = 25$$

Las sumas son todas cuadrados perfectos: $1^2, 2^2, 3^2, 4^2, 5^2$. Parece razonable conjeturar que este patrón seguirá sosteniéndose; esto es, la suma de números impares consecutivos, a partir de 1, siempre será un cuadrado perfecto. Intente algo más práctico. Si la suma es n^2, entonces el último número impar en la suma es $2n - 1$. (Compruebe esto en los cinco casos anteriores.) En símbolos, la conjetura se convierte en



$$1 + 3 + 5 + \cdots + (2n - 1) = n^2 \quad \text{para toda} \quad n \geq 1 \tag{1}$$

Note que la fórmula (1) en realidad es una colección *infinita* de enunciados, uno para cada valor de $n \geq 1$. Aunque la conjetura parece razonable, no es posible suponer que el patrón continúa, es necesario demostrarlo. Es aquí donde interviene la ***inducción matemática***.

Primer principio de inducción matemática

Sea $S(n)$ un enunciado acerca del entero positivo n. Si

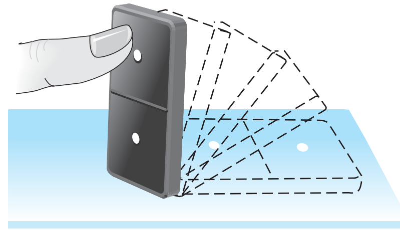
1. $S(1)$ es verdadero y
2. para toda $k \geq 1$, la verdad de $S(k)$ implica la verdad de $S(k + 1)$

entonces $S(n)$ es verdadero para toda $n \geq 1$.

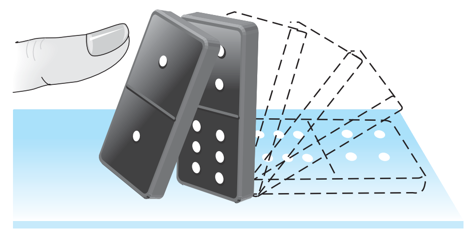
Verificar que $S(1)$ es verdadero se llama ***paso base***. La suposición de que $S(k)$ es verdadero para alguna $k \geq 1$ se le llama ***hipótesis de inducción***. Al uso de la hipótesis de inducción para probar que $S(k + 1)$ es entonces verdadera se le llama ***paso de inducción***. A la inducción matemática se le refiere como el *principio dominó* porque es análoga a demostrar que una línea de fichas de dominó caerá si (1) la primera ficha puede derribarse (el paso base) y (2) derribar cualquier ficha (la hipótesis de inducción) derribará la siguiente ficha (el paso de inducción). Vea la figura B.1.

Ahora se usa el principio de inducción matemática para demostrar la fórmula (1).

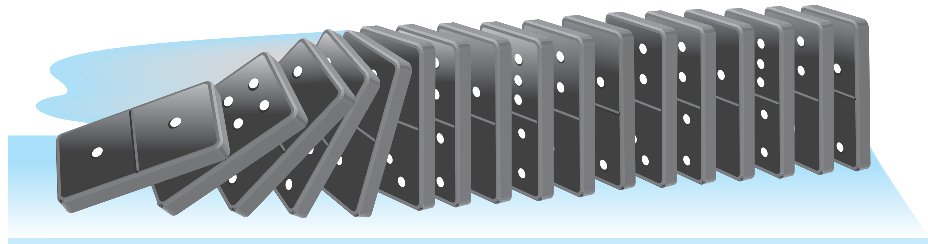
*En el sitio Web para el estudiante puede encontrar ejercicios y respuestas seleccionadas con número impar para este apéndice.



Si la primera ficha cae, y . . .



cada ficha que cae derriba a la siguiente . . .



entonces puede hacerse caer a todas las fichas al empujar la primera.

Figura B.1

Ejemplo B.1

Use inducción matemática para demostrar que

$$1 + 3 + 5 + \cdots + (2n - 1) = n^2$$

para toda $n \geq 1$.

Solución Para $n = 1$, la suma en el lado izquierdo es 1, mientras que el lado derecho es 1^2. Puesto que $1 = 1^2$, esto completa el paso base.

Ahora suponga que la fórmula es verdadera para algún entero $k \geq 1$. Esto es, suponga que

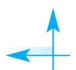
$$1 + 3 + 5 + \cdots + (2k - 1) = k^2$$

(Esta es la hipótesis de inducción.) El paso de inducción consiste en demostrar que la fórmula es verdadera cuando $n = k + 1$. Se ve que, cuando $n = k + 1$, el lado izquierdo de la fórmula (1) es

$$\begin{aligned} 1 + 3 + 5 + \cdots + (2(k+1) - 1) &= 1 + 3 + 5 + \cdots + (2k + 1) \\ &= \underbrace{1 + 3 + 5 + \cdots + (2k - 1)} + (2k + 1) \\ &= \qquad\qquad k^2 \qquad\qquad + 2k + 1 \quad \leftarrow \text{por la hipótesis de inducción} \\ &= (k + 1)^2 \end{aligned}$$

que es el lado derecho de la fórmula (1) cuando $n = k + 1$.

Esto completa el paso de inducción, y se concluye que la fórmula (1) es verdadera para toda $n \geq 1$, por el principio de inducción matemática.



El siguiente ejemplo brinda una demostración de una fórmula útil para la suma de los primeros n enteros positivos. La fórmula aparece varias veces en el texto; por ejemplo, vea la solución al ejercicio 51 de la sección 2.4.

Ejemplo B.2 Demuestre que

$$1 + 2 + \cdots + n = \frac{n(n+1)}{2}$$

para toda $n \geq 1$.

Solución La fórmula es verdadera para $n = 1$, pues

$$1 = \frac{1(1+1)}{2}$$

Suponga que la fórmula es verdadera para $n = k$; esto es,

$$1 + 2 + \cdots + k = \frac{k(k+1)}{2}$$

Es necesario demostrar que la fórmula es verdadera cuando $n = k + 1$; esto es, debe demostrar que

$$1 + 2 + \cdots + (k+1) = \frac{(k+1)[(k+1)+1]}{2}$$

Pero se ve que

$$\begin{aligned}
1 + 2 + \cdots + (k+1) &= (1 + 2 + \cdots + k) + (k+1) \\
&= \frac{k(k+1)}{2} + (k+1) \quad \text{para la hipótesis de inducción} \\
&= \frac{k(k+1) + 2(k+1)}{2} \\
&= \frac{k^2 + 3k + 2}{2} \\
&= \frac{(k+1)(k+2)}{2} \\
&= \frac{(k+1)[(k+1)+1]}{2}
\end{aligned}$$

que es lo que se necesita demostrar.

Esto completa el paso de inducción y se concluye que la fórmula es verdadera para toda $n \geq 1$, por el principio de inducción matemática.

En un modo similar, puede demostrar que la suma de los cuadrados de los primeros n enteros positivos satisface la fórmula

$$1^2 + 2^2 + 3^2 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

para toda $n \geq 1$. (Verifique esto.)

El paso base no necesita ser para $n = 1$, como ilustran los dos ejemplos siguientes.

Ejemplo B.3

Demuestre que $n! > 2^n$ para todos los enteros $n \geq 4$.

Solución El paso base aquí es cuando $n = 4$. Claramente, la desigualdad es verdadera en este caso, pues

$$4! = 24 > 16 = 2^4$$

Suponga que $k! > 2^k$ para algún entero $k \geq 4$. Entonces

$$\begin{aligned}(k+1)! &= (k+1)k! \\ &> (k+1)2^k && \text{por la hipótesis de inducción} \\ &\geq 5 \cdot 2^k && \text{pues } k \geq 4 \\ &> 2 \cdot 2^k = 2^{k+1}\end{aligned}$$

lo que verifica la desigualdad para $n = k + 1$ y completa el paso de inducción.

Se concluye que $n! > 2^n$ para todos los enteros $n \geq 4$, por el principio de inducción matemática.

Si a es un número real distinto de cero y $n \geq 0$ es un entero, puede ofrecer una definición recurrente de la potencia a^n que sea compatible con la inducción matemática. Se define $a^0 = 1$ y, para $n \geq 0$,

$$a^{n+1} = a^n a$$

(Esta forma evita los puntos suspensivos que se usan en la versión $a^n = \overbrace{aa \cdots a}^{n \text{ veces}}$.) Ahora puede usar inducción matemática para verificar una propiedad familiar de los exponentes.

Ejemplo B.4

Sea a un número real distinto de cero. Demuestre que $a^m a^n = a^{m+n}$ para todos los enteros $m, n \geq 0$.

Solución A primera vista, no está claro cómo proceder, pues existen *dos* variables, m y n. Pero simplemente es necesario tener una de ellas fija y realizar la inducción usando la otra. De este modo, sea $m \geq 0$ un entero fijo. Cuando $n = 0$, se tiene

$$a^m a^0 = a^m \cdot 1 = a^m = a^{m+0}$$

al usar la definición $a^0 = 1$. Por tanto, el paso base es verdadero.

Ahora suponga que la fórmula se cumple cuando $n = k$, donde $k \geq 0$. Entonces, $a^m a^k = a^{m+k}$. Para $n = k + 1$, use la definición recursiva y el hecho de que la suma y la multiplicación son asociativas, para ver que

$$\begin{aligned}a^m a^{k+1} &= a^m(a^k a) && \text{por definición} \\ &= (a^m a^k)a \\ &= a^{m+k}a && \text{por la hipótesis de inducción} \\ &= a^{(m+k)+1} && \text{por definición} \\ &= a^{m+(k+1)}\end{aligned}$$

Por tanto, la fórmula es verdadera para $n = k + 1$, y el paso de inducción está completo.

Se concluye que $a^m a^n = a^{m+n}$ para todos los enteros $m, n \geq 0$, por el principio de inducción matemática.

En los ejemplos B.1 al B.4, el uso de la hipótesis de inducción durante el paso de inducción es relativamente directo. Sin embargo, este no siempre es el caso. Con frecuencia es más útil una versión alternativa del principio de inducción matemática.

Segundo principio de inducción matemática

Sea $S(n)$ un enunciado acerca del entero positivo n. Si

1. $S(1)$ es verdadero y
2. la veracidad de $S(1), S(2), \ldots, S(k)$ implica la veracidad de $S(k + 1)$

entonces $S(n)$ es verdadero para toda $n \geq 1$.

La única diferencia entre los dos principios de inducción matemática está en la hipótesis de inducción: la primera versión supone que $S(k)$ es verdadera, mientras que la segunda versión supone que todos los $S(1), S(2), \ldots, S(k)$ son verdaderos. Esto hace parecer al segundo principio más débil que el primero, pues es necesario suponer más con la finalidad de demostrar $S(k+1)$ (aunque, paradójicamente, el segundo principio en ocasiones se llama de inducción *fuerte*). Sin embargo, de hecho, los dos principios son lógicamente equivalentes: cada uno implica al otro. (¿Puede ver por qué?)

El siguiente ejemplo presenta un caso donde el segundo principio de inducción matemática es más sencillo de usar que el primero. Recuerde que un número primo es un entero positivo cuyos únicos factores enteros positivos son 1 y él mismo.

Ejemplo B.5 Demuestre que todo entero positivo $n \geq 2$ es primo o puede factorizarse en un producto de primos.

Solución Claramente el resultado es verdadero cuando $n = 2$, pues 2 es primo. Ahora suponga que para todos los enteros n entre 2 y k, n es primo o puede factorizarse en un producto de primos. Sea $n = k + 1$. Si $k + 1$ es primo, ya terminó. De otro modo, debe factorizarlo en un producto de dos enteros más pequeños, por decir,

$$k + 1 = ab$$

Puesto que $2 \leq a, b \leq k$ (¿por qué?), la hipótesis de inducción se aplica a a y b. Por tanto,

$$a = p_1 \cdots p_r \quad \text{y} \quad b = q_1 \cdots q_s$$

donde los p y q son todos primos. Entonces

$$ab = p_1 \cdots p_r q_1 \cdots q_s$$

produce una factorización de ab en primos, lo que completa el paso de inducción.

Se concluye que el resultado es verdadero para todos los enteros $n \geq 2$, por el segundo principio de inducción matemática.



¿Ve por qué el primer principio de inducción matemática hubiera sido difícil de usar aquí? Se concluye con un ejemplo enormemente no trivial que involucra una combinación de inducción e inducción *hacia atrás*. El resultado es la desigualdad media aritmética-media geométrica que se estudió en el capítulo 7 en la Exploración: "Desigualdades geométricas y problemas de optimización". La astuta demostración del ejemplo B.6 se debe a Cauchy.

Ejemplo B.6 Sean $x_1, \ldots, x_n$ números reales no negativos. Demuestre que

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}$$

para todos los enteros $n \geq 2$.

Solución Para $n = 2$, la desigualdad se convierte en $\sqrt{xy} \leq (x + y)/2$. En los problemas 1 y 2 de la Exploración mencionada anteriormente se le pide verificar esto.

Si $S(n)$ es la desigualdad establecida, se probará que $S(k)$ implica $S(2k)$. Suponga que $S(k)$ es verdadero; esto es,

$$\sqrt[k]{x_1 x_2 \cdots x_k} \leq \frac{x_1 + x_2 + \cdots + x_k}{k}$$

para todo número real no negativo $x_1, \ldots, x_k$. Sean

$$x_1 = \frac{y_1 + y_2}{2}, \quad x_2 = \frac{y_3 + y_4}{2}, \quad \ldots, \quad x_k = \frac{y_{2k-1} + y_{2k}}{2}$$

Entonces

$$\begin{aligned}
\sqrt[2k]{y_1 \cdots y_{2k}} &= \sqrt[k]{\sqrt{y_1 \cdots y_{2k}}} = \sqrt[k]{\sqrt{y_1 y_2} \cdots \sqrt{y_{2k-1} y_{2k}}} \\
&\leq \sqrt[k]{\left(\frac{y_1 + y_2}{2}\right) \cdots \left(\frac{y_{2k-1} + y_{2k}}{2}\right)} && \text{por } S(2) \\
&= \sqrt[k]{x_1 \cdots x_k} \\
&\leq \frac{x_1 + x_2 + \cdots + x_k}{k} && \text{por } S(k) \\
&= \frac{\left(\frac{y_1 + y_2}{2}\right) + \cdots + \left(\frac{y_{2k-1} + y_{2k}}{2}\right)}{k} \\
&= \frac{y_1 + \cdots + y_{2k}}{2k}
\end{aligned}$$

lo que verifica $S(2k)$.

Por tanto, la desigualdad media aritmética-media geométrica es verdadera para $n = 2, 4, 8, \ldots$ las potencias de 2. Con la finalidad de completar la demostración, es necesario "rellenar los huecos". Se usará inducción hacia atrás para demostrar que $S(k)$ implica $S(k - 1)$. Si supone que $S(k)$ es verdadero, sea

$$x_k = \frac{x_1 + x_2 + \cdots + x_{k-1}}{k - 1}$$

Entonces

$$\sqrt[k]{x_1x_2\cdots x_{k-1}\left(\frac{x_1+x_2+\cdots+x_{k-1}}{k-1}\right)} \le \frac{x_1+x_2+\cdots+\left(\dfrac{x_1+x_2+\cdots+x_{k-1}}{k-1}\right)}{k}$$

$$= \frac{kx_1+kx_2+\cdots+kx_{k-1}}{k(k-1)}$$

$$= \frac{x_1+x_2+\cdots+x_{k-1}}{k-1}$$

De manera equivalente,

$$x_1x_2\cdots x_{k-1}\left(\frac{x_1+x_2+\cdots+x_{k-1}}{k-1}\right) \le \left(\frac{x_1+x_2+\cdots+x_{k-1}}{k-1}\right)^k$$

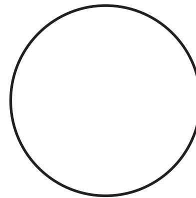
o
$$x_1x_2\cdots x_{k-1} \le \left(\frac{x_1+x_2+\cdots+x_{k-1}}{k-1}\right)^{k-1}$$

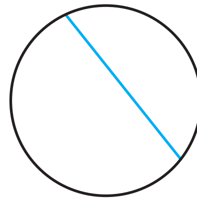
Al tomar la $(k-1)$-ésima raíz de ambos lados se obtiene $S(k-1)$.

Las dos inducciones, tomadas en conjunto, demuestran que la desigualdad media aritmética-media geométrica es verdadera para toda $n \ge 2$.



Comentario Aunque la inducción matemática es una herramienta poderosa e indispensable, no puede hacer milagros. Esto es, no puede demostrar que un patrón o fórmula se cumple si no lo hace. Considere los diagramas en la figura B.2, que muestran el número máximo de regiones $R(n)$ en las que se puede subdividir un círculo mediante n líneas rectas.


$R(0) = 1 = 2^0$


$R(1) = 2 = 2^1$

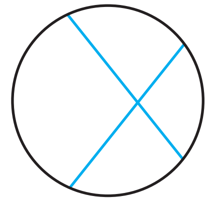

$R(2) = 4 = 2^2$

Figura B.2

Con base en la evidencia de la figura B.2, puede conjeturar que $R(n) = 2^n$ para $n \ge 0$ y tratar de demostrar esta conjetura usando inducción matemática. No tendrá éxito ¡porque esta fórmula no es correcta! Si hubiera considerado un caso más, habría descubierto que $R(3) = 7 \ne 8 = 2^3$, lo que por tanto derrumba la conjetura. De hecho, la fórmula correcta evidencia ser

$$R(n) = \frac{n^2+n+2}{2}$$



que *puede* verificarse mediante inducción. (¿Puede hacerlo?)

Para otros ejemplos en los que un patrón parezca ser verdadero, sólo para desaparecer cuando se consideran suficientes casos, vea el delicioso artículo de Richard K. Guy, "The Strong Law of Small Numbers" en el *American Mathematical Monthly*, Vol. 95 (1988), pp. 697-712.

Apéndice C*

Números complejos

[La] extensión del concepto número para incluir los irracionales, y de una vez agregaremos los imaginarios, es el mayor paso hacia adelante que ha dado jamás la matemática pura.

—Hermann Hankel
Theorie der Complexen Zahlensysteme
Leipzig, 1867, p. 60

Un ***número complejo*** tiene la forma $a + bi$, donde a y b son números reales e i es un símbolo con la propiedad de que $i^2 = -1$. El número real a se considera un tipo especial de número complejo, pues $a = a + 0i$. Si $z = a + bi$ es un número complejo, entonces la ***parte real*** de z, denotada Re z, es a, y la ***parte imaginaria*** de z, denotada Im z, es b. Dos números complejos $a + bi$ y $c + di$ son ***iguales*** si sus partes reales son iguales y sus partes imaginarias son iguales; esto es, si $a = c$ y $b = d$. Un número complejo $a + bi$ puede identificarse con el punto (a, b) y graficarse en el plano (llamado ***plano complejo***, o ***plano de Argand***), como se muestra en la figura C.1. En el plano complejo, el eje horizontal se llama ***eje real*** y el eje vertical se llama ***eje imaginario.***

No hay nada de "imaginario" en los números complejos, son tan "reales" como los números reales. El término *imaginario* surgió del estudio de ecuaciones polinomiales como $x^2 + 1 = 0$, cuyas soluciones no son "reales" (es decir, números reales). Vale la pena recordar que en alguna época los números negativos también se consideraron "imaginarios".

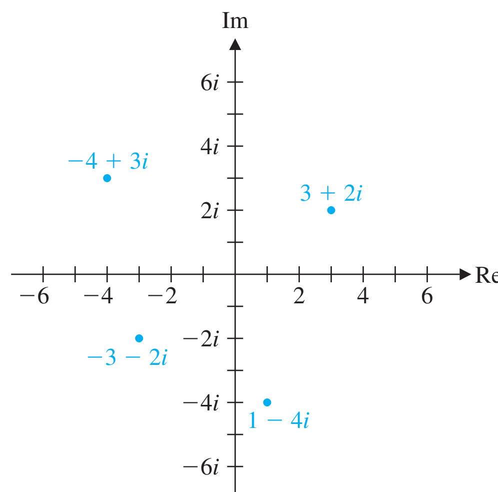


Figura C.1
El plano complejo

Jean-Robert Argand (1768-1822) fue un contador y matemático aficionado francés. Su interpretación geométrica de los números complejos apareció en 1806 en un libro que publicó de manera privada. Sin embargo, no fue el primero en brindar tal interpretación. El topógrafo noruego-danés Caspar Wessel (1745-1818) proporcionó la misma versión del plano complejo en 1787, pero su artículo no fue apreciado por la comunidad matemática sino hasta después de su muerte.

Operaciones con números complejos

La ***suma*** de los números complejos $a + bi$ y $c + di$ se define como

$$(a + bi) + (c + di) = (a + c) + (b + d)i$$

Note que, con la identificación de $a + bi$ con (a, b), $c + di$ con (c, d) y $(a + c) + (b + d)i$ con $(a + c, b + d)$, la suma de números complejos es igual que la suma vectorial. El ***producto*** de $a + bi$ y $c + di$ es

$$\begin{aligned}(a + bi)(c + di) &= a(c + di) + bi(c + di)\\ &= ac + adi + bci + bdi^2\end{aligned}$$

*En el sitio Web para el estudiante puede encontrar ejercicios y respuestas seleccionadas con número impar para este apéndice.

Dado que $i^2 = -1$, esta expresión se simplifica a $(ac - bd) + (ad + bc)i$. Por tanto, se tiene

$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i$$

Observe que, como caso especial, $a(c + di) = ac + adi$, de modo que el ***negativo*** de $c + di$ es $-(c + di) = (-1)(c + di) = -c - di$. Este hecho permite calcular la ***diferencia*** de $a + bi$ y $c + di$ como

$$\begin{aligned}(a + bi) - (c + di) &= (a + bi) + (-1)(c + di)\\ &= (a + (-c)) + (b + (-d))i\\ &= (a - c) + (b - d)i\end{aligned}$$

Ejemplo C.1

Encuentre la suma, diferencia y producto de $3 - 4i$ y $-1 + 2i$.

Solución La suma es

$$(3 - 4i) + (-1 + 2i) = (3 - 1) + (-4 + 2)i = 2 - 2i$$

La diferencia es

$$(3 - 4i) - (-1 + 2i) = (3 - (-1)) + (-4 - 2)i = 4 - 6i$$

El producto es

$$\begin{aligned}(3 - 4i)(-1 + 2i) &= -3 + 6i + 4i - 8i^2\\ &= -3 + 10i - 8(-1) = 5 + 10i\end{aligned}$$

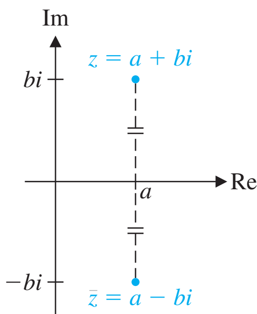


Figura C.2
Conjugados complejos

El ***conjugado*** de $z = a + bi$ es el número complejo

$$\bar{z} = a - bi$$

($\bar{z}$ se pronuncia "*z* barra".) La figura C.2 brinda la interpretación geométrica del conjugado.

Para encontrar el cociente de dos números complejos, multiplique el numerador y el denominador por el conjugado del denominador.

Ejemplo C.2

Exprese $\dfrac{-1 + 2i}{3 + 4i}$ en la forma $a + bi$.

Solución Multiplique el numerador y el denominador por $\overline{3 + 4i} = 3 - 4i$. Al usar el ejemplo C.1 se obtiene

$$\frac{-1 + 2i}{3 + 4i} = \frac{-1 + 2i}{3 + 4i} \cdot \frac{3 - 4i}{3 - 4i} = \frac{5 + 10i}{3^2 + 4^2} = \frac{5 + 10i}{25} = \frac{1}{5} + \frac{2}{5}i$$

A continuación se presenta un resumen de las propiedades de los conjugados. Las demostraciones se obtienen de la definición de conjugado; debe verificarlas usted mismo.

1. $\bar{\bar{z}} = z$
2. $\overline{z + w} = \bar{z} + \bar{w}$
3. $\overline{zw} = \bar{z}\bar{w}$
4. Si $z \neq 0$, entonces $\overline{\left(\dfrac{w}{z}\right)} = \dfrac{\bar{w}}{\bar{z}}$
5. z es real si y sólo si $\bar{z} = z$.

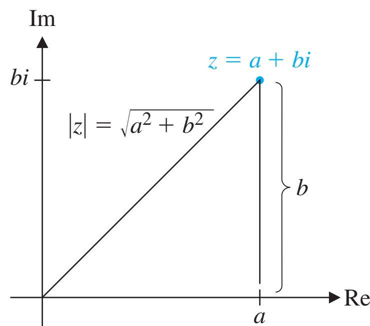


Figura C.3

El ***valor absoluto*** (o ***módulo***) $|z|$ de un número complejo $z = a + bi$ es su distancia desde el origen. Como muestra la figura C.3, el teorema de Pitágoras produce

$$|z| = |a + bi| = \sqrt{a^2 + b^2}$$

Observe que

$$z\bar{z} = (a + bi)(a - bi) = a^2 - abi + bai - b^2i^2 = a^2 + b^2$$

Por tanto,
$$z\bar{z} = |z|^2$$

Esto proporciona una forma alternativa de describir el proceso de división para el cociente de dos números complejos. Si w y $z \neq 0$ son dos números complejos, entonces

$$\frac{w}{z} = \frac{w}{z} \cdot \frac{\bar{z}}{\bar{z}} = \frac{w\bar{z}}{z\bar{z}} = \frac{w\bar{z}}{|z|^2}$$



A continuación se presenta un resumen de algunas de las propiedades del valor absoluto. (Debe tratar de demostrarlas usando la definición de valor absoluto y otras propiedades de los números complejos.)

1. $|z| = 0$ si y sólo si $z = 0$.
2. $|z| = |\bar{z}|$
3. $|zw| = |z||w|$
4. Si $z \neq 0$, entonces $\left|\dfrac{1}{z}\right| = \dfrac{1}{|z|}$.
5. $|z + w| \leq |z| + |w|$

Forma polar

Como ya vio, el número complejo $z = a + bi$ puede representarse geométricamente por el punto (a, b). Este punto también puede expresarse en términos de ***coordenadas polares*** (r, θ), donde $r \geq 0$, como se muestra en la figura C.4. Se tiene

$$a = r\cos\theta \quad \text{y} \quad b = r\,\text{sen}\,\theta$$

de modo que
$$z = a + bi = r\cos\theta + (r\,\text{sen}\,\theta)i$$

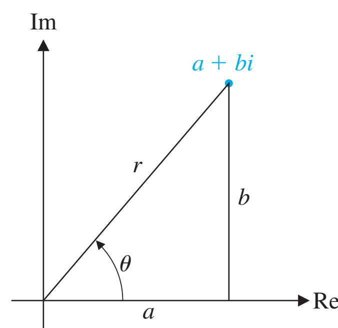


Figura C.4

Por tanto, cualquier número complejo puede escribirse en la ***forma polar***

$$z = r(\cos\theta + i\,\text{sen}\,\theta)$$

donde $r = |z| = \sqrt{a^2 + b^2}$ y $\tan\theta = b/a$. El ángulo θ se llama ***argumento*** de z y se denota arg z. Observe que arg z no es único: sumar o restar cualquier múltiplo entero de 2π produce otro argumento de z. Sin embargo, sólo hay un argumento θ que satisface

$$-\pi < \theta \leq \pi$$

A este se le llama ***argumento principal*** de z y se denota Arg z.

Ejemplo C.3

Escriba los siguientes números complejos en forma polar usando sus argumentos principales:

(a) $z = 1 + i$ (b) $w = 1 - \sqrt{3}i$

Solución (a) Calcule

$$r = |z| = \sqrt{1^2 + 1^2} = \sqrt{2} \quad \text{y} \quad \tan\theta = \frac{1}{1} = 1$$

Por tanto, Arg $z = \theta = \dfrac{\pi}{4}$ $(= 45°)$, y se tiene

$$z = \sqrt{2}\left(\cos\frac{\pi}{4} + i\,\text{sen}\,\frac{\pi}{4}\right)$$

como se muestra en la figura C.5.

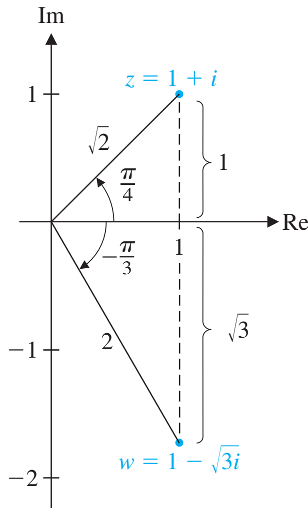


Figura C.5

(b) Se tiene

$$r = |w| = \sqrt{1^2 + (-\sqrt{3})^2} = \sqrt{4} = 2 \quad \text{y} \quad \tan\theta = \frac{-\sqrt{3}}{1} = -\sqrt{3}$$

Puesto que w está en el cuarto cuadrante, debe tener Arg $z = \theta = -\dfrac{\pi}{3}$ $(= -60°)$. Por tanto,

$$w = 2\left(\cos\left(-\frac{\pi}{3}\right) + i\,\text{sen}\left(-\frac{\pi}{3}\right)\right)$$

Vea la figura C.5.

La forma polar de los números complejos se puede usar para ofrecer interpretaciones geométricas de multiplicación y división. Sean

$$z_1 = r_1(\cos\theta_1 + i\,\text{sen}\,\theta_1) \quad \text{y} \quad z_2 = r_2(\cos\theta_2 + i\,\text{sen}\,\theta_2)$$

Al multiplicar, se obtiene

$$\begin{aligned} z_1 z_2 &= r_1 r_2(\cos\theta_1 + i\,\text{sen}\,\theta_1)(\cos\theta_2 + i\,\text{sen}\,\theta_2) \\ &= r_1 r_2[(\cos\theta_1\cos\theta_2 - \text{sen}\,\theta_1\,\text{sen}\,\theta_2) + i(\text{sen}\,\theta_1\cos\theta_2 + \cos\theta_1\,\text{sen}\,\theta_2)] \end{aligned}$$

Al usar identidades trigonométricas

$$\begin{aligned} \cos(\theta_1 + \theta_2) &= \cos\theta_1\cos\theta_2 - \text{sen}\,\theta_1\,\text{sen}\,\theta_2 \\ \text{sen}(\theta_1 + \theta_2) &= \text{sen}\,\theta_1\cos\theta_2 + \cos\theta_1\,\text{sen}\,\theta_2 \end{aligned}$$

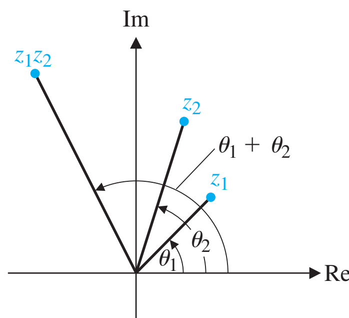


Figura C.6

se obtiene

$$z_1z_2 = r_1r_2[\cos(\theta_1 + \theta_2) + i\,\text{sen}(\theta_1 + \theta_2)] \tag{1}$$

que es la forma polar de un número complejo con valor absoluto r_1r_2 y argumento $\theta_1 + \theta_2$. Esto demuestra que

$$|z_1z_2| = |z_1||z_2| \quad \text{y} \quad \arg(z_1z_2) = \arg z_1 + \arg z_2$$

La fórmula (1) dice que *para multiplicar dos números complejos, se multiplican sus valores absolutos y se suman sus argumentos.* Vea la figura C.6.

De igual modo, al usar las identidades de resta para seno y coseno, es posible demostrar que

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}[\cos(\theta_1 - \theta_2) + i\,\text{sen}(\theta_1 - \theta_2)] \quad \text{si } z \neq 0$$



(Verifique esto.) Por tanto,

$$\left|\frac{z_1}{z_2}\right| = \frac{|z_1|}{|z_2|} \quad \text{y} \quad \arg\left(\frac{z_1}{z_2}\right) = \arg z_1 - \arg z_2$$

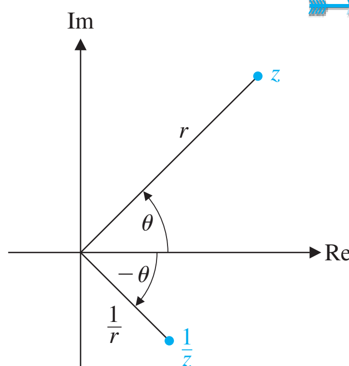


Figura C.7

y se ve que *para dividir dos números complejos se dividen sus valores absolutos y se restan sus argumentos.*

Como caso especial del último resultado, se obtiene una fórmula para el recíproco de un número complejo en forma polar. Al hacer $z_1 = 1$ (y por tanto $\theta_1 = 0$) y $z_2 = z$ (y por tanto $\theta_2 = \theta$), se obtiene lo siguiente

Si $z = r(\cos\theta + i\,\text{sen}\,\theta)$ es distinto de cero, entonces

$$\frac{1}{z} = \frac{1}{r}(\cos\theta - i\,\text{sen}\,\theta)$$

Vea la figura C.7.

Ejemplo C.4

Encuentre el producto de $1 + i$ y $1 - \sqrt{3}i$ en forma polar.

Solución A partir del ejemplo C.3 se tiene

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i\,\text{sen}\,\frac{\pi}{4}\right) \quad \text{y} \quad 1 - \sqrt{3}i = 2\left(\cos\left(-\frac{\pi}{3}\right) + i\,\text{sen}\left(-\frac{\pi}{3}\right)\right)$$

Por tanto,

$$\begin{aligned}(1 + i)(1 - \sqrt{3}i) &= 2\sqrt{2}\left[\cos\left(\frac{\pi}{4} - \frac{\pi}{3}\right) + i\,\text{sen}\left(\frac{\pi}{4} - \frac{\pi}{3}\right)\right] \\ &= 2\sqrt{2}\left[\cos\left(-\frac{\pi}{12}\right) + i\,\text{sen}\left(-\frac{\pi}{12}\right)\right]\end{aligned}$$

Vea la figura C.8.

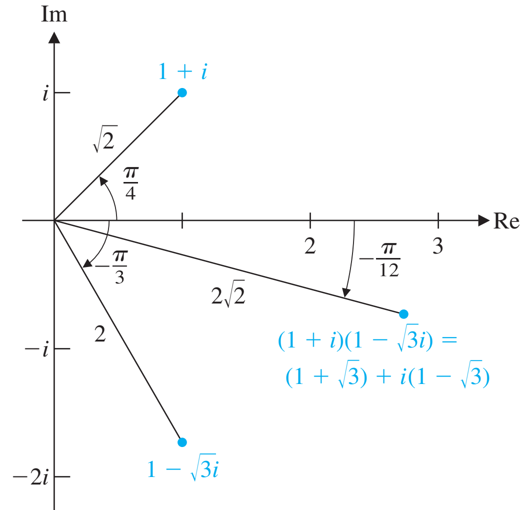
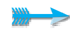


Figura C.8



Comentario Dado que $(1 + i)(1 - \sqrt{3}i) = (1 + \sqrt{3}) + i(1 - \sqrt{3})$(compruébelo), debe tener

$$1 + \sqrt{3} = 2\sqrt{2}\cos\left(-\frac{\pi}{12}\right) = -2\sqrt{2}\cos\left(\frac{\pi}{12}\right)$$

y

$$1 - \sqrt{3} = 2\sqrt{2}\,\text{sen}\left(-\frac{\pi}{12}\right) = -2\sqrt{2}\,\text{sen}\left(\frac{\pi}{12}\right)$$



(¿Por qué?) Esto implica que

$$\cos\left(\frac{\pi}{12}\right) = \frac{1 + \sqrt{3}}{2\sqrt{2}} \quad \text{y} \quad \text{sen}\left(\frac{\pi}{12}\right) = \frac{\sqrt{3} - 1}{2\sqrt{2}}$$

Por tanto, se tiene un método para encontrar el seno y el coseno de un ángulo como $\pi/12$, que no es un ángulo especial, sino que puede obtenerse como una suma o diferencia de ángulos especiales.

Teorema de De Moivre

Si n es un entero positivo y $z = r(\cos\theta + i\,\text{sen}\,\theta)$, entonces el uso repetido de la fórmula (1) produce fórmulas para las potencias de z:

$$z^2 = r^2(\cos 2\theta + i\,\text{sen}\,2\theta)$$
$$z^3 = zz^2 = r^3(\cos 3\theta + i\,\text{sen}\,3\theta)$$
$$z^4 = zz^3 = r^4(\cos 4\theta + i\,\text{sen}\,4\theta)$$
$$\vdots$$

Abraham De Moivre (1667-1754) fue un matemático francés que realizó importantes aportaciones a la trigonometría, geometría analítica, probabilidad y estadística.

En general, se tiene el siguiente resultado, conocido como ***teorema de De Moivre***.

Teorema C.1 **Teorema de De Moivre**

Si $z = r(\cos\theta + i \operatorname{sen}\theta)$ y n es un entero positivo, entonces

$$z^n = r^n(\cos n\theta + i \operatorname{sen} n\theta)$$

Dicho de manera diferente, se tiene

$$|z^n| = |z|^n \quad \text{y} \quad \arg(z^n) = n \arg z$$

En palabras, el teorema de De Moivre dice que *para calcular la n-ésima potencia de un número complejo, se evalúa la n-ésima potencia de su valor absoluto y su argumento se multiplica por n.*

Ejemplo C.5 Encuentre $(1 + i)^6$.

Solución A partir del ejemplo C.3(a) se tiene

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i \operatorname{sen}\frac{\pi}{4}\right)$$

Por tanto, el teorema de De Moivre produce

$$\begin{aligned}(1 + i)^6 &= (\sqrt{2})^6\left(\cos\frac{6\pi}{4} + i \operatorname{sen}\frac{6\pi}{4}\right)\\ &= 8\left(\cos\frac{3\pi}{2} + i \operatorname{sen}\frac{3\pi}{2}\right)\\ &= 8(0 + i(-1)) = -8i\end{aligned}$$

Vea la figura C.9 que muestra $1 + i, (1 + i)^2, (1 + i)^3, \ldots, (1 + i)^6$.

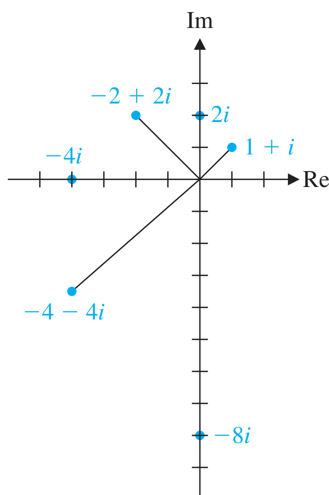


Figura C.9
Potencias de $1 + i$

También es posible usar el teorema de De Moivre para encontrar raíces n-ésimas de números complejos. Una raíz n-ésima del número complejo z es cualquier número complejo w tal que

$$w^n = z$$

En forma polar se tiene

$$w = s(\cos\varphi + i \operatorname{sen}\varphi) \quad \text{y} \quad z = r(\cos\theta + i \operatorname{sen}\theta)$$

de modo que, por el teorema de De Moivre,

$$s^n(\cos n\varphi + i \operatorname{sen} n\varphi) = r(\cos\theta + i \operatorname{sen}\theta)$$

Al igualar los valores absolutos, se ve que

$$s^n = r \quad \text{o} \quad s = r^{1/n} = \sqrt[n]{r}$$

También se debe tener

$$\cos n\varphi = \cos\theta \quad \text{y} \quad \operatorname{sen} n\varphi = \operatorname{sen}\theta$$



(¿Por qué?) Puesto que las funciones seno y coseno tienen cada una periodo 2π, dichas ecuaciones implican que $n\varphi$ y θ difieren por un múltiplo entero de 2π; esto es,

$$n\varphi = \theta + 2k\pi \quad \text{o} \quad \varphi = \frac{\theta + 2k\pi}{n}$$

donde k es un entero. Por tanto,

$$w = r^{1/n}\left[\cos\left(\frac{\theta + 2k\pi}{n}\right) + i \operatorname{sen}\left(\frac{\theta + 2k\pi}{n}\right)\right]$$

describe las posibles raíces n-ésimas de z conforme k varía sobre los enteros. No es difícil demostrar que $k = 0, 1, 2, \ldots, n - 1$ produce distintos valores de w, de modo que existen exactamente n diferentes raíces n-ésimas de $z = r(\cos\theta + i \operatorname{sen}\theta)$. Este resultado se resume del modo siguiente:

Sea $z = r(\cos\theta + i \operatorname{sen}\theta)$ y sea n un entero positivo. Entonces z tiene exactamente n distintas raíces n-ésimas dadas por

$$r^{1/n}\left[\cos\left(\frac{\theta + 2k\pi}{n}\right) + i \operatorname{sen}\left(\frac{\theta + 2k\pi}{n}\right)\right] \tag{2}$$

para $k = 0, 1, 2, \ldots, n - 1$.

Ejemplo C.6

Encuentre las tres raíces cúbicas de -27.

Solución En forma polar, $-27 = 27(\cos\pi + i \operatorname{sen}\pi)$. Se tiene que las raíces cúbicas de -27 están dadas por

$$(-27)^{1/3} = 27^{1/3}\left[\cos\left(\frac{\pi + 2k\pi}{3}\right) + i \operatorname{sen}\left(\frac{\pi + 2k\pi}{3}\right)\right] \quad \text{para} \quad k = 0, 1, 2$$

Al usar la fórmula (2) con $n = 3$, se obtiene

$$27^{1/3}\left[\cos\frac{\pi}{3} + i \operatorname{sen}\frac{\pi}{3}\right] = 3\left(\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) = \frac{3}{2} + \frac{3\sqrt{3}}{2}i$$

$$27^{1/3}\left[\cos\left(\frac{\pi + 2\pi}{3}\right) + i \operatorname{sen}\left(\frac{\pi + 2\pi}{3}\right)\right] = 3(\cos\pi + i \operatorname{sen}\pi) = -3$$

$$27^{1/3}\left[\cos\left(\frac{\pi + 4\pi}{3}\right) + i \operatorname{sen}\left(\frac{\pi + 4\pi}{3}\right)\right] = 3\left(\cos\frac{5\pi}{3} + i \operatorname{sen}\frac{5\pi}{3}\right)$$

$$= 3\left(\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) = \frac{3}{2} - \frac{3\sqrt{3}}{2}i$$

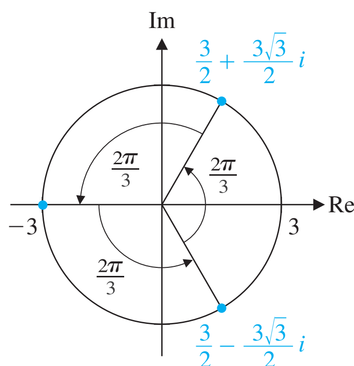


Figura C.10
Las raíces cúbicas de -27

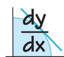
Como muestra la figura C.10, las tres raíces cúbicas de -27 están igualmente espaciadas $2\pi/3$ radianes (120°) alrededor de un círculo de radio 3 con centro en el origen.

En general, la fórmula (2) implica que las raíces n-ésimas de $z = r(\cos\theta + i \operatorname{sen}\theta)$ se hallaran en un círculo de radio $r^{1/n}$ con centro en el origen. Más aún, estarán igualmente espaciados $2\pi/n$ radianes $(360/n°)$. (Verifique esto.) Por ende, si puede encontrar una raíz n-ésima de z, las restantes raíces n-ésimas de z pueden obtenerse al rotar la primera raíz en incrementos sucesivos de $2\pi/n$ radianes. De haber sabido esto en el ejemplo C.6, podría haber usado el hecho de que la raíz cúbica real de -27 es -3 y entonces rotarla dos veces un ángulo de $2\pi/3$ radianes (120°) para obtener las otras dos raíces cúbicas.

Leonhard Euler (1707-1783) fue el más prolífico matemático de todos los tiempos. Tiene más de 900 publicaciones con su nombre y sus obras completas llenan más de 70 volúmenes. Existen tantos resultados atribuidos a él que "fórmula de Euler" o "teorema de Euler" puede significar muchas cosas diferentes, dependiendo del contexto.

Euler trabajó en tantas áreas de la matemáticas que es difícil mencionarlas todas. Sus aportaciones al cálculo y al análisis, a las ecuaciones diferenciales, la teoría de números, la geometría, la topología, la mecánica y otras áreas de las matemáticas aplicadas siguen siendo influyentes. También introdujo gran parte de la notación que se usa en la actualidad, incluyendo π, e, i, Σ para suma, Δ para diferencia y $f(x)$ para una función, y fue el primero en tratar a seno y coseno como funciones.

Euler nació en Suiza, pero pasó gran parte de su vida matemática en Rusia y Alemania. En 1727 se unió a la Academia de Ciencias de San Petersburgo, que fue fundada por Catalina I, esposa de Pedro El Grande. Viajó a Berlín en 1741 a invitación de Federico el Grande, pero regresó a San Petersburgo en 1766, donde permaneció hasta su muerte. Cuando era joven, perdió la visión en un ojo como resultado de una enfermedad, y hacia 1776 perdió la visión en el otro ojo y quedó totalmente ciego. Notablemente, sus resultados matemáticos no disminuyeron y siguió siendo productivo hasta el día de su muerte.



Fórmula de Euler

En cálculo se aprende que la función e^z tiene una expansión en serie de potencias

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots$$

que converge para todo número real z. Puede demostrarse que esta expansión también funciona cuando z es un número complejo y que la función exponencial compleja e^z obedece las reglas usuales para exponentes. Las funciones seno y coseno también tienen expansiones en series de potencias:

$$\operatorname{sen} x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + - \cdots$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + - \cdots$$

Si se hace $z = ix$, donde x es un número real, entonces se tiene

$$e^z = e^{ix} = 1 + ix + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \cdots$$

Al usar el hecho de que $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, $i^5 = i$, etcétera, que se repite en un ciclo de longitud 4, se ve que

$$\begin{aligned} e^{ix} &= 1 + ix - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \frac{ix^5}{5!} - \frac{x^6}{6!} - \frac{ix^7}{7!} + + - - \cdots \\ &= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + - \cdots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + - \cdots\right) \\ &= \cos x + i \operatorname{sen} x \end{aligned}$$

Este notable resultado se conoce como ***fórmula de Euler***.

Teorema C.2 **Fórmula de Euler**

Para cualquier número real x,

$$e^{ix} = \cos x + i \operatorname{sen} x$$

Al usar la fórmula de Euler, se ve que la forma polar de un número complejo puede escribirse en forma más compacta como

$$z = r(\cos\theta + i \operatorname{sen}\theta) = re^{i\theta}$$

Por ejemplo, del ejemplo C.3(a), se tiene

$$1 + i = \sqrt{2}\left(\cos\frac{\pi}{4} + i \operatorname{sen}\frac{\pi}{4}\right) = \sqrt{2}e^{i\pi/4}$$

También es posible ir en la dirección contraria y convertir un exponencial complejo de vuelta en forma polar o estándar.

Ejemplo C.7 Escriba las siguientes en la forma $a + bi$:

(a) $e^{i\pi}$ (b) $e^{2+i\pi/4}$

Solución (a) Al usar la fórmula de Euler se tiene

$$e^{i\pi} = \cos\pi + i \operatorname{sen}\pi = -1 + i \cdot 0 = -1$$

(Si esta ecuación se escribe como $e^{i\pi} + 1 = 0$, se obtiene la que seguramente es una de las ecuaciones más notables de las matemáticas. Contiene las operaciones fundamentales de suma, multiplicación y exponenciación; la identidad aditiva 0 y la identidad multiplicativa 1; los dos números trascendentales más importantes, π y e; y la unidad compleja i... ¡todo en una ecuación!)

(b) Al usar las reglas para exponentes junto con la fórmula de Euler se obtiene

$$e^{2+i\pi/4} = e^2 e^{i\pi/4} = e^2\left(\cos\frac{\pi}{4} + i \operatorname{sen}\frac{\pi}{4}\right) = e^2\left(\frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}\right)$$

$$= \frac{e^2\sqrt{2}}{2} + \frac{e^2\sqrt{2}}{2}i$$

Si $z = re^{i\theta} = r(\cos\theta + i \operatorname{sen}\theta)$, entonces

$$\bar{z} = r(\cos\theta - i \operatorname{sen}\theta) \tag{3}$$

Las identidades trigonométricas

$$\cos(-\theta) = \cos\theta \text{ y } \operatorname{sen}(-\theta) = -\operatorname{sen}\theta$$

permiten reescribir la ecuación (3) como

$$\bar{z} = r(\cos(-\theta) + i\,\text{sen}(-\theta)) = re^{i(-\theta)}$$

Esto produce la siguiente fórmula útil para el conjugado:

Si $z = re^{i\theta}$, entonces

$$\bar{z} = re^{-i\theta}$$

Nota La fórmula de Euler proporciona una rápida demostración en una línea del teorema de De Moivre:

$$[r(\cos\theta + i\,\text{sen}\,\theta)]^n = (re^{i\theta})^n = r^n e^{in\theta} = r^n(\cos n\theta + i\,\text{sen}\,n\theta)$$

Apéndice D*

Polinomios

Un ***polinomio*** es una función p de una sola variable x que puede escribirse en la forma

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \quad (1)$$

Euler ofreció la más algebraica de las demostraciones de la existencia de las raíces de una ecuación [polinominal]. . . Considero que es injusto acreditar esta demostración exclusivamente a Gauss, quien simplemente agregó los toques finales.

—George Frobenius, 1907
citado en el archivo MacTutor de Historia de las Matemáticas, http://www-history.mcs.st-and.ac.uk/history/

donde $a_0, a_1, \ldots, a_n$ son constantes ($a_n \neq 0$), llamadas ***coeficientes*** de p. Con la convención de que $x^0 = 1$, puede usar la notación suma para escribir p como

$$p(x) = \sum_{k=0}^{n} a_kx^k$$

El entero n se llama ***grado*** de p, que se denota al escribir grado $p = n$. Un polinomio de grado cero se llama ***polinomio constante***.

Ejemplo D.1

¿Cuáles de los siguientes son polinomios?

(a) $2 - \frac{1}{3}x + \sqrt{2}x^2$ (b) $2 - \dfrac{1}{3x^2}$ (c) $\sqrt{2x^2}$

(d) $\ln\left(\dfrac{2e^{5x^3}}{e^{3x}}\right)$ (e) $\dfrac{x^2 - 5x + 6}{x - 2}$ (f) $\sqrt{x}$

(g) $\cos(2\cos^{-1}x)$ (h) e^x

Solución (a) Este es el único que obviamente es un polinomio.

(b) Un polinomio de la forma que se muestra en la ecuación (1) no puede volverse infinito conforme x tiende a un valor finito $[\lim_{x \to c} p(x) \neq \pm\infty]$, mientras que $2 - 1/3x^2$ tiende a $-\infty$ conforme x tiende a cero. Por tanto, no es un polinomio.

(c) Se tiene

$$\sqrt{2x^2} = \sqrt{2}\sqrt{x^2} = \sqrt{2}|x|$$

que es igual a $\sqrt{2}x$ cuando $x \geq 0$ y a $-\sqrt{2}x$ cuando $x < 0$. Por tanto, esta expresión se forma al "empalmar" dos polinomios (un *polinomio en trozos*), pero no es un polinomio en sí mismo.

*En el sitio Web para el estudiante puede encontrar ejercicios y respuestas seleccionadas con número impar para este apéndice.

(d) Al usar las propiedades de los exponentes y los logaritmos se tiene

$$\ln\left(\frac{2e^{5x^3}}{e^{3x}}\right) = \ln(2e^{5x^3-3x}) = \ln 2 + \ln(e^{5x^3-3x})$$
$$= \ln 2 + 5x^3 - 3x = \ln 2 - 3x + 5x^3$$

de modo que esta expresión es un polinomio.

(e) El dominio de esta función consiste de todos los números reales $x \neq 2$. Para dichos valores de x, la función se simplifica a

$$\frac{x^2 - 5x + 6}{x - 2} = \frac{(x-2)(x-3)}{x-2} = x - 3$$

así que es posible decir que es un polinomio *sobre su dominio.*

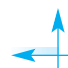
(f) Se ve que esta función no puede ser un polinomio (incluso sobre su dominio $x \geq 0$), pues la derivación repetida de un polinomio de la forma que se muestra en la ecuación (1) eventualmente resultará en cero y $\sqrt{x}$ no tiene esta propiedad. (Verifique esto.)

(g) El dominio de esta expresión es $-1 \leq x \leq 1$. Sea $\theta = \cos^{-1} x$ de modo que $\cos\theta = x$. Al usar una identidad trigonométrica, se ve que

$$\cos(2\cos^{-1} x) = \cos 2\theta = 2\cos^2\theta - 1 = 2x^2 - 1$$

de modo que esta expresión es un polinomio sobre su dominio.

(h) Al analizar esta expresión como se hizo en (f), se concluye que no es un polinomio.



Dos polinomios son ***iguales*** si los coeficientes de las correspondientes potencias de x son todas iguales. En particular, los polinomios iguales deben tener el mismo grado. La ***suma*** de dos polinomios se obtiene al sumar los coeficientes de las correspondientes potencias de x.

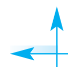
Ejemplo D.2

Encuentre la suma de $2 - 4x + x^2$ y $1 + 2x - x^2 + 3x^3$.

Solución Calcule

$$(2 - 4x + x^2) + (1 + 2x - x^2 + 3x^3) = (2 + 1) + (-4 + 2)x$$
$$+ (1 + (-1))x^2 + (0 + 3)x^3$$
$$= 3 - 2x + 3x^3$$

donde el primer polinomio se "rellenó" al darle un coeficiente de cero a una x^3.



La ***diferencia*** de dos polinomios se define de manera análoga, al restar coeficientes en lugar de sumarlos. El ***producto*** de dos polinomios se obtiene al usar repetidamente la ley distributiva y luego reunir las correspondientes potencias de x.

Ejemplo D.3

Encuentre el producto de $2 - 4x + x^2$ y $1 + 2x - x^2 + 3x^3$.

Solución Se obtiene

$$\begin{aligned}
&(2 - 4x + x^2)(1 + 2x - x^2 + 3x^3)\\
&\quad = 2(1 + 2x - x^2 + 3x^3) - 4x(1 + 2x - x^2 + 3x^3)\\
&\qquad + x^2(1 + 2x - x^2 + 3x^3)\\
&\quad = (2 + 4x - 2x^2 + 6x^3) + (-4x - 8x^2 + 4x^3 - 12x^4)\\
&\qquad + (x^2 + 2x^3 - x^4 + 3x^5)\\
&\quad = 2 + (4x - 4x) + (-2x^2 - 8x^2 + x^2) + (6x^3 + 4x^3 + 2x^3)\\
&\qquad + (-12x^4 - x^4) + 3x^5\\
&\quad = 2 - 9x^2 + 12x^3 - 13x^4 + 3x^5
\end{aligned}$$

Observe que para dos polinomios p y q se tiene

$$\text{grado}(pq) = \text{grado } p + \text{grado } q$$

Si p y q son polinomios con grado $q \leq$ grado p, puede dividir q entre p y usar división larga para obtener el cociente p/q. El siguiente ejemplo ilustra el procedimiento, que es igual para la división larga de un entero entre otro. Así como el cociente de dos enteros no es, en general, un entero, el cociente de dos polinomios no es, en general, otro polinomio.

Ejemplo D.4

Calcule $\dfrac{1 + 2x - x^2 + 3x^3}{2 - 4x + x^2}$.

Solución Se realizará división larga. Es útil escribir cada polinomio con potencias *decrecientes* de x. En concordancia, se tiene

$$x^2 - 4x + 2 \,\overline{\big)\, 3x^3 - x^2 + 2x + 1}$$

Comience por dividir $3x^3$ entre x^2 para obtener el cociente parcial $3x$. Luego multiplique $3x$ por el divisor $x^2 - 4x + 2$, reste el resultado y baje el siguiente término del dividendo $(3x^3 - x^2 + 2x + 1)$:

$$\begin{array}{r l}
 & \quad\;\; 3x \\
x^2 - 4x + 2 \big) & \overline{3x^3 - x^2 + 2x + 1} \\
 & \underline{3x^3 - 12x^2 + 6x \phantom{{}+1}} \\
 & \qquad 11x^2 - 4x + 1
\end{array}$$

Luego repita el proceso con $11x^2$, multiplique 11 por $x^2 - 4x + 2$ y reste el resultado de $11x^2 - 4x + 1$. Se obtiene

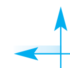
$$\begin{array}{r l}
 & \qquad\quad 3x + 11 \\
x^2 - 4x + 2 \big) & \overline{3x^3 - x^2 + 2x + 1} \\
 & \underline{3x^3 - 12x^2 + 6x \phantom{{}+22}} \\
 & \qquad 11x^2 - 4x + 1 \\
 & \qquad \underline{11x^2 - 44x + 22} \\
 & \qquad\qquad\quad 40x - 21
\end{array}$$

Ahora se tiene un residuo de $40x - 21$. Su grado es menor que el del divisor $x^2 - 4x + 2$, de modo que el proceso se detiene y se encontró que

$$3x^3 - x^2 + 2x + 1 = (x^2 - 4x + 2)(3x + 11) + (40x - 21)$$

o

$$\frac{3x^3 - x^2 + 2x + 1}{x^2 - 4x + 2} = 3x + 11 + \frac{40x - 21}{x^2 - 4x + 2}$$



El ejemplo D.4 puede generalizarse para producir el siguiente resultado, conocido como ***algoritmo de la división***.

Teorema D.1 **El algoritmo de la división**

Si f y g son polinomios con grado $g \leq$ grado f, entonces existen polinomios q y r tales que

$$f(x) = g(x)q(x) + r(x)$$

donde $r = 0$ o grado $r <$ grado g.

En el ejemplo D.4,

$$f(x) = 3x^3 - x^2 + 2x + 1, \quad g(x) = x^2 - 4x + 2, \quad q(x) = 3x + 11,$$

$$\text{y} \quad r(x) = 40x - 21$$

En el algoritmo de la división, si el residuo es cero, entonces

$$f(x) = g(x)q(x)$$

y se dice que g es un ***factor*** de f. (Note que q también es un factor de f.) Existe una estrecha conexión entre los factores de un polinomio y sus ceros. Un ***cero*** de un polinomio f es un número a tal que $f(a) = 0$. [El número a también se llama ***raíz*** de la ecuación polinomial $f(x) = 0$.] El siguiente resultado, conocido como ***teorema del Factor***, establece la conexión entre los factores de un polinomio y sus ceros.

Teorema D.2 **El teorema del factor**

Sea f un polinomio y sea a una constante. Entonces a es un cero de f si y sólo si $x - a$ es un factor de $f(x)$.

Demostración Por el algoritmo de la división,

$$f(x) = (x - a)q(x) + r(x)$$

donde $r(x) = 0$ o grado $r <$ grado $(x - a) = 1$. Por ende, en cualquier caso, $r(x) = r$ es una constante. Ahora,

$$f(a) = (a - a)q(a) + r = r$$

de modo que $f(a) = 0$ si y sólo si $r = 0$, lo cual es equivalente a

$$f(x) = (x - a)q(x)$$

como se quería probar.

No hay un método que esté garantizado para encontrar los ceros de un polinomio dado. Sin embargo, existen algunos lineamientos que son útiles en casos especiales. El caso de un polinomio con coeficientes *enteros* es particularmente interesante. El siguiente resultado, conocido como el ***teorema de raíces racionales***, brinda criterios para que un cero de tal polinomio sea un número *racional*.

Teorema D.3

El teorema de raíces racionales

Sea

$$f(x) = a_0 + a_1x + \cdots + a_nx^n$$

un polinomio con coeficientes enteros y sea a/b un número racional escrito en términos más bajos. Si a/b es un cero de f, entonces a_0 es un múltiplo de a y a_n es un múltiplo de b.

Demostración Si a/b es un cero de f, entonces

$$a_0 + a_1\left(\frac{a}{b}\right) + \cdots + a_{n-1}\left(\frac{a}{b}\right)^{n-1} + a_n\left(\frac{a}{b}\right)^n = 0$$

Al multiplicar por b^n, se tiene

$$a_0b^n + a_1ab^{n-1} + \cdots + a_{n-1}a^{n-1}b + a_na = 0 \qquad (1)$$

lo cual implica que

$$a_0b^n + a_1ab^{n-1} + \cdots + a_{n-1}a^{n-1}b = -a_na^n \qquad (2)$$

El lado izquierdo de la ecuación (2) es un múltiplo de b, de modo que a_na^n también debe ser un múltiplo de b. Puesto que a/b está en términos más bajos, a y b no tienen factores comunes mayores que 1. Por tanto, a_n debe ser un múltiplo de b.

La ecuación (1) también puede escribirse como

$$-a_0b^n = a_1ab^{n-1} + \cdots + a_{n-1}a^{n-1}b + a_na^n$$


y un argumento similar muestra que a_0 debe ser un múltiplo de a. (Demuestre esto.)

Ejemplo D.5

Encuentre todas las raíces racionales de la ecuación

$$6x^3 + 13x^2 - 4 = 0 \qquad (3)$$

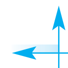
Solución Si a/b es una raíz de esta ecuación, entonces 6 es un múltiplo de b y -4 es un múltiplo de a, por el teorema de raíces racionales. Por tanto,

$$a \in \{\pm 1, \pm 2, \pm 4\} \qquad \text{y} \qquad b \in \{\pm 1, \pm 2, \pm 3, \pm 6\}$$

Al formar todos los posibles números racionales a/b con estas elecciones de a y b, se ve que las únicas raíces racionales posibles de la ecuación dada son

$$\pm 1, \pm 2, \pm 4, \pm\tfrac{1}{2}, \pm\tfrac{1}{3}, \pm\tfrac{2}{3}, \pm\tfrac{4}{3}, \pm\tfrac{1}{6}$$

Al sustituir dichos valores en la ecuación (3) uno a la vez, se encuentra que -2, $-\frac{2}{3}$ y $\frac{1}{2}$ son los únicos valores de esta lista que realmente son raíces. (Compruebe esto.) Como verá dentro de poco, una ecuación polinomial de grado 3 no puede tener más de tres raíces, de modo que estas no sólo son todas las raíces *racionales* de la ecuación (3) sino también sus *únicas* raíces.



El método de ensayo y error del ejemplo D.5 puede mejorarse en varias formas. Por ejemplo, una vez que encuentre una raíz a de una ecuación polinomial dada $f(x) = 0$, se sabe que $x - a$ es un factor de $f(x)$, por decir, $f(x) = (x - a)g(x)$. Por tanto, puede dividir $f(x)$ entre $x - a$ (mediante división larga) para encontrar $g(x)$. Puesto que grado $g <$ grado f, las raíces de $g(x) = 0$ [que también son raíces de $f(x) = 0$] pueden ser más fáciles de encontrar. En particular, si $g(x)$ es un polinomio cuadrático, se tiene acceso a la ***fórmula cuadrática***.

Suponga

$$ax^2 + bx + c = 0$$

(Puede suponer que a es positivo, pues multiplicar ambos lados por -1 produciría una ecuación equivalente de otro modo.) Entonces, al completar el cuadrado, se tiene

$$a\left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2}\right) = \frac{b^2}{4a} - c$$

(Verifique esto.) De manera equivalente,

$$a\left(x + \frac{b}{2a}\right)^2 = \frac{b^2}{4a} - c \quad \text{o} \quad \left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

Por tanto,

$$x + \frac{b}{2a} = \pm\sqrt{\frac{b^2 - 4ac}{4a^2}} = \frac{\pm\sqrt{b^2 - 4ac}}{2a}$$

o

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Vuelva a revisar la ecuación del ejemplo D.5 con la fórmula cuadrática en mente.

Ejemplo D.6

Encuentre las raíces de $6x^3 + 13x^2 - 4 = 0$.

Solución Suponga que usa el teorema de raíces racionales para descubrir que $x = -2$ es una raíz racional de $6x^3 + 13x^2 - 4 = 0$. Entonces $x + 2$ es un factor de $6x^3 + 13x^2 - 4$ y la división larga produce

$$6x^3 + 13x^2 - 4 = (x + 2)(6x^2 + x - 2)$$

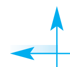

(Compruebe esto.) Ahora puede aplicar la fórmula cuadrática al segundo factor para encontrar que sus ceros son

$$x = \frac{-1 \pm \sqrt{1^2 - 4(6)(-2)}}{2 \cdot 6}$$

$$= \frac{-1 \pm \sqrt{49}}{12} = \frac{-1 \pm 7}{12}$$

$$= \tfrac{6}{12}, -\tfrac{8}{12}$$

o, en términos más bajos, $\frac{1}{2}$ y $-\frac{2}{3}$. Por tanto, las tres raíces de la ecuación (3) son -2, $\frac{1}{2}$, y $-\frac{2}{3}$, como se determinó en el ejemplo D.5.



Comentario El teorema del factor establece una conexión entre los ceros de un polinomio y sus factores *lineales*. Sin embargo, un polinomio sin factores lineales todavía puede tener factores de grado mayor. Más aún, cuando se pide factorizar un polinomio, es necesario conocer el sistema numérico al que se supone pertenecen los coeficientes de los factores.

Por ejemplo, considere el polinomio

$$p(x) = x^4 + 1$$

Sobre los *números racionales* $\mathbb{Q}$, los únicos ceros posibles de p son 1 y -1, por el teorema de raíces racionales. Una revisión rápida muestra que en realidad ninguno de ellos funciona, de modo que $p(x)$ no tiene factores *lineales* con coeficientes racionales, por el teorema del factor. Sin embargo, $p(x)$ todavía puede factorizarse en un producto de dos *cuadráticas*. Los factores cuadráticos se verificarán usando el método de ***coeficientes indeterminados***.

Suponga que

$$x^4 + 1 = (x^2 + ax + b)(x^2 + cx + d)$$


Al expandir el lado derecho y comparar coeficientes, se obtienen las ecuaciones

$$\begin{aligned} a + c &= 0 \\ b + ac + d &= 0 \\ bc + ad &= 0 \\ bd &= 1 \end{aligned}$$

Si $a = 0$, entonces $c = 0$ y $d = -b$. Esto produce $-b^2 = 1$, que no tiene soluciones en $\mathbb{Q}$. Por tanto, puede suponer que $a \neq 0$. Entonces $c = -a$ y se obtiene $d = b$. Ahora se tiene que $b^2 = 1$, de modo que $b = 1$ o $b = -1$. Esto implica que $a^2 = 2$ o $a^2 = -2$, respectivamente, ninguno de los cuales tiene solución en $\mathbb{Q}$. Se concluye que $x^4 + 1$ no puede factorizarse sobre $\mathbb{Q}$. Se dice que es ***irreductible*** sobre $\mathbb{Q}$.

Sin embargo, sobre los *números reales* $\mathbb{R}$, $x^4 + 1$ sí se factoriza. Los cálculos recién realizados demuestran que

$$x^4 + 1 = (x^2 + \sqrt{2}x + 1)(x^2 - \sqrt{2}x + 1)$$


(¿Por qué?) Para ver si es posible factorizar aún más, aplique la fórmula cuadrática. Se ve que el primer factor tiene ceros

$$x = \frac{-\sqrt{2} \pm \sqrt{(\sqrt{2})^2 - 4}}{2} = \frac{-\sqrt{2} \pm \sqrt{-2}}{2} = \frac{\sqrt{2}}{2}(-1 \pm i) = -\frac{1}{\sqrt{2}} \pm \frac{1}{\sqrt{2}}i$$

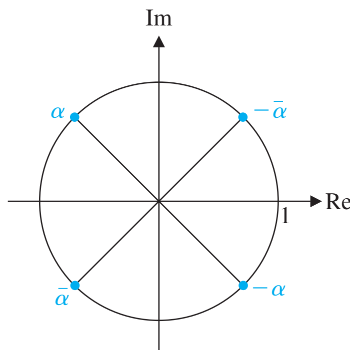


Figura D.1

que están en $\mathbb{C}$ mas no en $\mathbb{R}$. Por tanto, $x^2 + \sqrt{2}x + 1$ no puede factorizarse en factores lineales sobre $\mathbb{R}$. De igual modo, $x^2 - \sqrt{2}x + 1$ no puede factorizarse en factores lineales sobre $\mathbb{R}$.

Los cálculos demuestran que una factorización completa de $x^4 + 1$ es posible sobre los *números complejos* $\mathbb{C}$. Los cuatro ceros de $x^4 + 1$ son

$$\alpha = -\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i, \quad \bar{\alpha} = -\frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i, \quad -\bar{\alpha} = \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i,$$
$$-\alpha = \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}}i$$

que, como muestra la figura D.1, todas están en el círculo unitario en el plano complejo. Por tanto, la factorización de $x^4 + 1$ es

$$x^4 + 1 = (x - \alpha)(x - \bar{\alpha})(x + \bar{\alpha})(x + \alpha)$$

El comentario anterior ilustra muchas propiedades importantes de los polinomios. Note que el polinomio $p(x) = x^4 + 1$ satisface grado $p = 4$ y tiene exactamente cuatro ceros en $\mathbb{C}$. Más aún, sus ceros complejos ocurren en *pares conjugados*; esto es: sus ceros complejos pueden parearse como

$$\{\alpha, \bar{\alpha}\} \quad \text{y} \quad \{-\alpha, -\bar{\alpha}\}$$

Estos dos últimos hechos son verdaderos en general. El primero es un ejemplo del ***teorema fundamental del álgebra (TFA)***, un resultado que Gauss demostró por primera vez en 1797.

Teorema D.4 **El teorema fundamental del álgebra**

Todo polinomio de grado n con coeficientes reales o complejos tiene exactamente n ceros (contando multiplicidades) en $\mathbb{C}$.

Este importante teorema en ocasiones se enuncia como

"Todo polinomio con coeficientes reales o complejos tiene un cero en $\mathbb{C}$."

Llame a este enunciado TFA′. Ciertamente, TFA implica TFA′. Por el contrario, si TFA′ es verdadero, entonces, si se tiene un polinomio p de grado n, tiene un cero α en $\mathbb{C}$. El teorema del factor dice entonces que $x - \alpha$ es un factor de $p(x)$, de modo que

$$p(x) = (x - \alpha)q(x)$$

donde q es un polinomio de grado $n - 1$ (también con coeficientes reales o complejos). Ahora puede aplicar TFA' a q para obtener otro cero y así por el estilo, lo que hace verdadero a TFA. Este argumento puede convertirse en una agradable prueba de inducción.

(Inténtelo.)

No es posible dar una fórmula (a lo largo de las líneas de la fórmula cuadrática) para los ceros de polinomios de grado 5 o más. (El trabajo de Abel y Galois confirmó esto; vea la página 322.) En consecuencia, deben usarse otros métodos para probar TFA. La demostración que proporcionó Gauss usa métodos topológicos y puede encontrarse en cursos de matemáticas más avanzados.

Ahora suponga que

$$p(x) = a_0 + a_1x + \cdots + a_nx^n$$

es un polinomio con coeficientes reales. Sea α un cero complejo de p de modo que

$$a_0 + a_1\alpha + \cdots + a_n\alpha^n = p(\alpha) = 0$$

Entonces, al usar las propiedades de los conjugados, se tiene

$$\begin{aligned} p(\overline{\alpha}) &= a_0 + a_1\overline{\alpha} + \cdots + a_n\overline{\alpha}^n = \overline{a}_0 + \overline{a}_1\overline{\alpha} + \cdots + \overline{a}_n\overline{\alpha}^n \\ &= \overline{a_0 + a_1\alpha + \cdots + a_n\alpha^n} \\ &= \overline{p(\alpha)} = \overline{0} = 0 \end{aligned}$$

Por tanto, $\overline{\alpha}$ también es un cero de p. Esto demuestra el siguiente resultado:

Los ceros complejos de un polinomio con coeficientes reales se presentan en pares conjugados.

En algunas situaciones no es necesario saber *cuáles* son los ceros de un polinomio, sólo es necesario saber *dónde* se ubican. Por ejemplo, acaso sólo sea necesario conocer si los ceros son positivos o negativos (como en el Teorema 4.35). Un teorema que es útil en este aspecto es la ***regla de los signos de Descartes***. Ello permite hacer ciertas predicciones acerca del número de ceros positivos de un polinomio con coeficientes reales con base en los signos de dichos coeficientes.

Descartes enunció esta regla en su libro de 1637 *La Gèometrie*, pero no ofreció una demostración. Más tarde, varios matemáticos proporcionaron una demostración y Gauss ofreció una versión un poco más precisa del teorema en 1828.

Dado un polinomio $a_0 + a_1x + \cdots + a_nx^n$, escriba en orden sus coeficientes distintos de cero. Sustituya cada coeficiente positivo con un signo más y cada coeficiente negativo con un signo menos. Se dirá que el polinomio tiene k ***cambios de signo*** si existen k lugares donde los coeficientes cambian de signo. Por ejemplo, el polinomio $2 - 3x + 4x^3 + x^4 - 7x^5$ tiene el patrón de signos

$$\underbrace{+\ -}\underbrace{\ \ +}\ \underbrace{+\ -}$$

.

de modo que tiene tres cambios de signo, como se indicó.

Teorema D.5 **Regla de los signos de Descartes**

Sea p un polinomio con coeficientes reales que tiene k cambios de signo. Entonces el número de ceros positivos de p (contando multiplicidades) es cuando mucho k.

En palabras: la regla de los signos de Descartes dice que un polinomio real no puede tener más ceros positivos que cambios de signo.

Ejemplo D.7 Demuestre que el polinomio $p(x) = 4 + 2x^2 - 7x^4$ tiene exactamente un cero positivo.

Solución Los coeficientes de p tienen el patrón de signos $+ + -$, que sólo tiene un cambio de signo. De modo que, por la regla de los signos de Descartes, p tiene cuando mucho un cero positivo. Pero $p(0) = 4$ y $p(1) = -1$, de modo que hay un cero en alguna parte en el intervalo $(0, 1)$. En consecuencia, este es el único cero positivo de p.

También puede usar la regla de los signos de Descartes para dar una cota sobre el número de ceros *negativos* de un polinomio con coeficientes reales. Sea

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

y sea b un cero negativo de p. Entonces $b = -c$ para $c > 0$ y se tiene

$$\begin{aligned} 0 = p(b) &= a_0 + a_1b + a_2b^2 + \cdots + a_nb^n \\ &= a_0 - a_1c + a_2c^2 - + \cdots + (-1)^n a_nc^n \end{aligned}$$

Pero $$p(-x) = a_0 - a_1x + a_2x^2 - + \cdots + (-1)^n a_nx^n$$

de modo que c es un cero positivo de $p(-x)$. Por tanto, $p(x)$ tiene exactamente tantos ceros negativos como $p(-x)$ tiene ceros positivos. Combinada con la regla de los signos de Descartes, esta observación produce lo siguiente:

Sea p un polinomio con coeficientes reales. Entonces el número de ceros negativos de p es cuando mucho el número de cambios de signo de $p(-x)$.

Ejemplo D.8

Demuestre que los ceros de $p(x) = 1 + 3x + 2x^2 + x^5$ no pueden ser todos reales.

Solución Los coeficientes de $p(x)$ no tienen cambios de signo, de modo que p no tiene ceros positivos. Puesto que $p(-x) = 1 - 3x + 2x^2 - x^5$ tiene tres cambios de signo entre sus coeficientes, p tiene cuando mucho tres ceros negativos. Note que 0 tampoco es un cero de p, de modo que p tiene cuando mucho tres ceros reales. Por tanto, tiene al menos dos ceros complejos (no reales).

Respuestas a ejercicios impares seleccionados

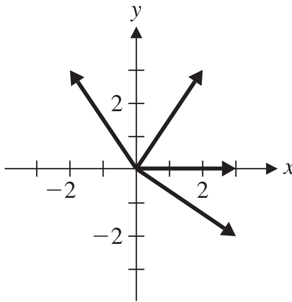
Las respuestas son sencillas. Plantear las preguntas correctas [es lo que es] difícil

—Doctor Who
"The Face of Evil,"
Por Chris Boucher
BBC, 1977

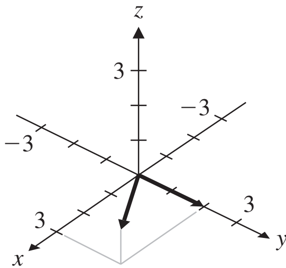
Capítulo 1

Ejercicios 1.1

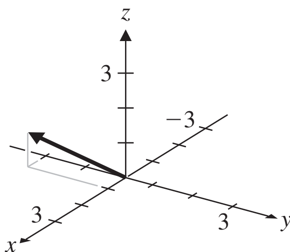
1.



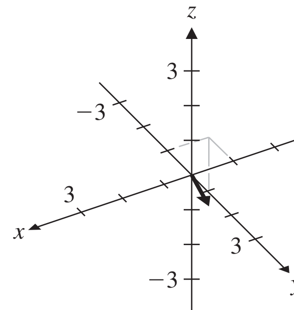
3. **(a)**, **(b)**



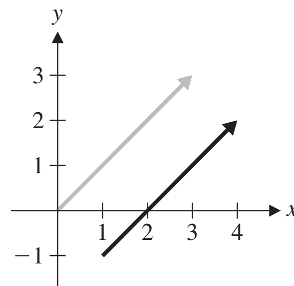
(c)



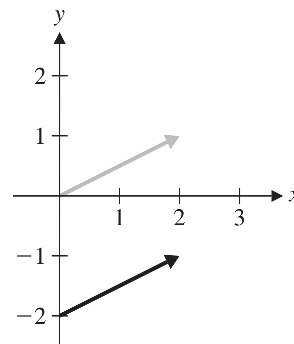
(d)



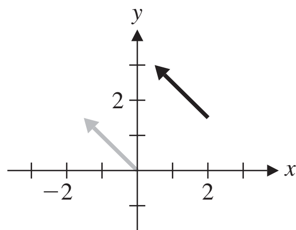
5. **(a)**



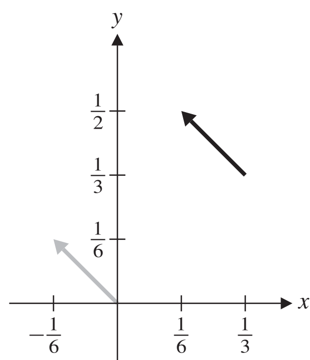
(b)



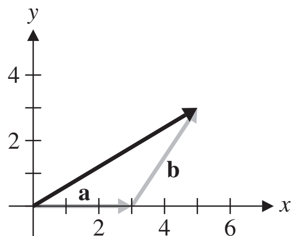
(c)



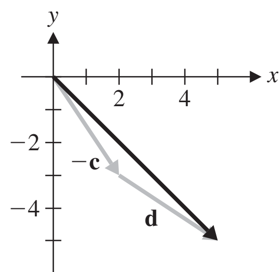
(d)



7. $\mathbf{a} + \mathbf{b} = [5, 3]$



9. $\mathbf{d} - \mathbf{c} = [5, -5]$



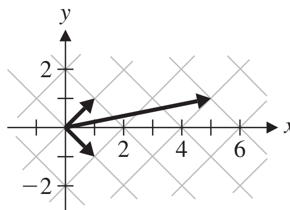
11. $[3, -2, 3]$

13. $\mathbf{u} = \begin{bmatrix} 1/2 \\ \sqrt{3}/2 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} -\sqrt{3}/2 \\ -1/2 \end{bmatrix}, \mathbf{u} + \mathbf{v} = \begin{bmatrix} (1 - \sqrt{3})/2 \\ (\sqrt{3} - 1)/2 \end{bmatrix}, \mathbf{u} - \mathbf{v} = \begin{bmatrix} (1 + \sqrt{3})/2 \\ (1 + \sqrt{3})/2 \end{bmatrix}$

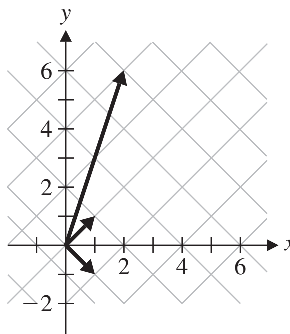
15. $5\mathbf{a}$

17. $\mathbf{x} = 3\mathbf{a}$

19.



21. $\mathbf{w} = -2\mathbf{u} + 4\mathbf{v}$



25. $\mathbf{u} + \mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{u} \cdot \mathbf{v} = 1$

27. $\mathbf{u} + \mathbf{v} = [0, 1, 0, 0], \mathbf{u} \cdot \mathbf{v} = 1$

29.

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

31. 0 **33.** 1

35. 0 **37.** 2, 0, 3

39. 5 **41.** [1, 1, 0]

43. 3, 2 **45.** $x = 2$

47. No hay solución **49.** $x = 3$

51. No hay solución **53.** $x = 2$

55. $x = 1$ o $x = 5$

57. **(a)** Toda $a \neq 0$ **(b)** $a = 1, 5$
(c) a y m no pueden tener factores comunes distintos de 1 [esto es, el máximo común divisor (mcd) de a y m es 1.]

Ejercicios 1.2

1. -1 **3.** 11 **5.** 2

7. $\sqrt{5}, \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ **9.** $\sqrt{14}, \begin{bmatrix} 1/\sqrt{14} \\ 2/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$

11. $\sqrt{6}, [1/\sqrt{6}, 1/\sqrt{3}, 1/\sqrt{2}, 0]$

13. $\sqrt{17}$ **15.** $\sqrt{6}$

17. **(a)** $\mathbf{u} \cdot \mathbf{v}$ es un escalar, no un vector.

(c) $\mathbf{v} \cdot \mathbf{w}$ es un escalar y $\mathbf{u}$ es un vector.

19. Agudo **21.** Agudo **23.** Agudo

25. 60° **27.** ≈88.10° **29.** ≈14.34°

31. Puesto que $\overrightarrow{AB} \cdot \overrightarrow{AC} = \begin{bmatrix} -4 \\ 1 \\ -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ -3 \end{bmatrix} = 0$, $\angle BAC$ es un ángulo recto.

33. Si se toma el cubo como un cubo unitario (como en la figura 1.34), las cuatro diagonales están dadas por los vectores

$$\mathbf{d}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{d}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \mathbf{d}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \mathbf{d}_4 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

Dado que $\mathbf{d}_i \cdot \mathbf{d}_j \neq 0$ para toda $i \neq j$ (seis posibilidades), ningún par de diagonales es perpendicular.

35. $D = (-2, 1, 1)$

37. 5 mi/h a un ángulo de ≈ 53.13° hacia la orilla

39. 60°

41. $\begin{bmatrix} -\frac{3}{5} \\ \frac{4}{5} \end{bmatrix}$ **43.** $\begin{bmatrix} \frac{3}{2} \\ -\frac{3}{2} \\ \frac{3}{2} \\ -\frac{3}{2} \end{bmatrix}$ **45.** $\begin{bmatrix} -0.301 \\ 0.033 \\ -0.252 \end{bmatrix}$

47. $\mathcal{A} = \sqrt{45}/2$ **49.** $k = -2, 3$

51. $\mathbf{v}$ es de la forma $k\begin{bmatrix} b \\ -a \end{bmatrix}$, donde k es un escalar.

53. Se violaría la desigualdad de Cauchy-Schwarz.

Ejercicios 1.3

1. **(a)** $\begin{bmatrix} 3 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = 0$ **(b)** $3x + 2y = 0$

3. **(a)** $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 3 \end{bmatrix}$

(b) $x = 1 - t$
$y = 3t$

5. **(a)** $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ **(b)** $x = t$
$y = -t$
$z = 4t$

7. **(a)** $\begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 2$ **(b)** $3x + 2y + z = 2$

9. **(a)** $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = s\begin{bmatrix} 2 \\ 1 \\ 2 \end{bmatrix} + t\begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$

(b) $x = 2s - 3t$
$y = s + 2t$
$z = 2s + t$

11. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + t\begin{bmatrix} 2 \\ 2 \end{bmatrix}$

13. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + s\begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ -2 \end{bmatrix}$

15. **(a)** $x = t$
$y = -1 + 3t$
$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 3 \end{bmatrix}$

17. Los vectores directores para las dos recta están dados por $\mathbf{d}_1 = \begin{bmatrix} 1 \\ m_1 \end{bmatrix}$ y $\mathbf{d}_2 = \begin{bmatrix} 1 \\ m_2 \end{bmatrix}$. Las rectas son perpendiculares si y sólo si $\mathbf{d}_1$ y $\mathbf{d}_2$ son ortogonales. Pero $\mathbf{d}_1 \cdot \mathbf{d}_2 = 0$ si y sólo si $1 + m_1 m_2 = 0$ o, de manera equivalente $m_1 m_2 = -1$.

19. **(a)** Perpendicular **(b)** Paralelo
(c) Perpendicular **(d)** Perpendicular

21. $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} + t\begin{bmatrix} 3 \\ 2 \end{bmatrix}$

23. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 3 \end{bmatrix} + t\begin{bmatrix} -1 \\ 3 \\ -1 \end{bmatrix}$

25. **(a)** $x = 0, x = 1, y = 0, y = 1, z = 0, z = 1$
(b) $x - y = 0$ **(c)** $x + y - z = 0$

27. $3\sqrt{2}/2$ **29.** $2\sqrt{3}/3$ **31.** $(\frac{1}{2}, \frac{1}{2})$

33. $(\frac{4}{3}, \frac{4}{3}, \frac{8}{3})$ **35.** $18\sqrt{13}/13$ **37.** $\frac{5}{3}$

43. ≈78.9° **45.** ≈80.4°

Ejercicios 1.4

1. 13 N a aproximadamente N 67.38 E

3. $8\sqrt{3}$ N a un ángulo de 30° hacia $\mathbf{f}_1$

5. 4 N a un ángulo de 60° hacia $\mathbf{f}_2$

7. 5 N a un ángulo de 60° a la fuerza dada, $5\sqrt{3}$ N perpendicular hacia la fuerza de 5 N.

9. $750\sqrt{2}$ N

11. 980 N

13. ≈ 117.6 N en el alambre de 15 cm, ≈ 88.2 N en el alambre de 20 cm

15. [1, 0, 1, 1, 1]

17. No hay error

19. No hay error

21. $d = 2$

23. $d = 0$

27. $d = 0$

31. $d = X$

33. (b) $[0, 4, 4, 9, 9, 0, 8, 3, 5, 6]$

35. (b) $[0, 3, 8, 7, 0, 9, 9, 0, 2, 6]$

Preguntas de repaso

1. **(a)** V **(c)** F **(e)** V **(g)** F **(i)** V

3. $\mathbf{x} = \begin{bmatrix} 8 \\ 11 \end{bmatrix}$ **5.** 120° **7.** $\begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix}$

9. $2x + 3y - z = 7$ **11.** $\sqrt{6}/2$

13. Se violaría la desigualdad de Cauchy-Schwarz

15. $2\sqrt{6}/3$ **17.** $x = 2$ **19.** 3

Capítulo 2

Ejercicios 2.1

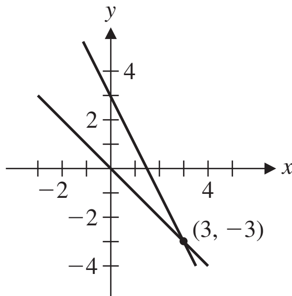
1. Lineal **3.** No es lineal debido al término x^{-1}

5. Lineal **7.** $2x + 4y = 7$

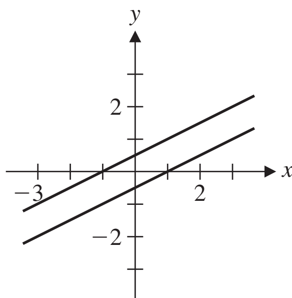
9. $x + y = 4 (x, y \neq 0)$

11. $\left\{ \begin{bmatrix} 2t \\ t \end{bmatrix} \right\}$ **13.** $\left\{ \begin{bmatrix} 4 - 2s - 3t \\ s \\ t \end{bmatrix} \right\}$

15. Solución única, $x = 3, y = -3$



17. No hay solución



19. $[7, 3]$ **21.** $[\frac{2}{3}, \frac{1}{3}, -\frac{1}{3}]$

23. $[5, -2, 1, 1]$ **25.** $[2, -7, -32]$

27. $\left[\begin{array}{rr|r} 1 & -1 & 0 \\ 2 & 1 & 3 \end{array}\right]$ **29.** $\left[\begin{array}{rr|r} 1 & 5 & -1 \\ -1 & 1 & -5 \\ 2 & 4 & 4 \end{array}\right]$

31. $\begin{aligned} y + z &= 1 \\ x - y \quad &= 1 \\ 2x - y + z &= 1 \end{aligned}$ **33.** $[1, 1]$

35. $[4, -1]$ **37.** No hay solución

39. **(a)** $\begin{aligned} 2x + \ \ y &= 3 \\ 4x + 2y &= 6 \end{aligned}$ **(b)** $\begin{aligned} x &= \tfrac{3}{2} - \tfrac{1}{2}s \\ y &= s \end{aligned}$

41. Sean $u = \dfrac{1}{x}$ y $v = \dfrac{1}{y}$. La solución es $x = \frac{1}{3}, y = -\frac{1}{2}$.

43. Sean $u = \tan x, v = \operatorname{sen} y, w = \cos z$. Una solución es $x = \pi/4, y = -\pi/6, z = \pi/3$. (Existe un número infinito de soluciones.)

Ejercicios 2.2

1. No **3.** Forma escalonada reducida por renglón

5. No **7.** No

9. **(a)** $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$ **11.** **(b)** $\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

13. **(b)** $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

15. Realice operaciones elementales con renglón en el orden $R_4 + 29R_3, 8R_3, R_4 - 3R_2, R_2 \leftrightarrow R_3, R_4 - R_1, R_3 + 2R_1$, y, finalmente, $R_2 + 2R_1$.

17. Una posibilidad es realizar operaciones elementales con renglón sobre A en el orden $R_2 - 3R_1, \frac{1}{2}R_2, R_1 + 2R_2, R_2 + 3R_1, R_1 \leftrightarrow R_2$.

19. *Sugerencia:* escoja una matriz aleatoria de 2×2 e intente esto, ¡cuidadosamente!

21. En realidad estas son dos operaciones elementales con renglón combinadas: $3R_2$ y $R_2 - 2R_1$.

23. Ejercicio 1: 3; ejercicio 3: 2; ejercicio 5: 2; ejercicio 7: 3

25. $\begin{bmatrix} 2 \\ 5 \\ 1 \end{bmatrix}$ **27.** $t\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ **29.** $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$

31. $\begin{bmatrix} 24 \\ -10 \\ 0 \\ 0 \\ 0 \end{bmatrix} + r\begin{bmatrix} 6 \\ -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + s\begin{bmatrix} 0 \\ 6 \\ 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} 12 \\ -6 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

33. No hay solución

35. Solución única

37. Número infinito de soluciones

39. *Sugerencia:* demuestre que si $ad - bc \neq 0$, el rank de $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ es 2. (Existen dos casos: $a = 0$ y $a \neq 0$.) Use el teorema del rank para deducir que el sistema dado debe tener una solución única.

41. **(a)** No hay solución si $k = -1$
(b) Una solución única si $k \neq \pm 1$
(c) Número infinito de soluciones si $k = 1$

43. **(a)** No hay solución si $k = 1$
(b) Una solución única si $k \neq -2, 1$
(c) Número infinito de soluciones si $k = -2$

45. $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} + t\begin{bmatrix} 9 \\ -10 \\ -7 \end{bmatrix}$

49. No hay intersección

51. Los vectores requeridos $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ son las soluciones del sistema homogéneo con matriz aumentada

$$\left[\begin{array}{ccc|c} u_1 & u_2 & u_3 & 0 \\ v_1 & v_2 & v_3 & 0 \end{array}\right]$$

Por el Teorema 3, existe un número infinito de soluciones. Si $u_1 \neq 0$ y $u_1v_2 - u_2v_1 \neq 0$, las soluciones están dadas por

$$t\begin{bmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{bmatrix}$$

Pero una comprobación directa demuestra que éstas todavía son soluciones incluso si $u_1 = 0$ y/o $u_1v_2 - u_2v_1 = 0$.

53. $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$ **55.** $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ **57.** $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$

Ejercicios 2.3

1. Sí **3.** No **5.** Sí **7.** Sí

9. Es necesario demostrar que la ecuación vectorial $x\begin{bmatrix} 1 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$ tiene una solución para todos los valores de a y b. Esta ecuación vectorial es equivalente al sistema lineal cuya matriz aumentada es

$\left[\begin{array}{cc|c} 1 & 1 & a \\ 1 & -1 & b \end{array}\right]$. La reducción por renglón produce

$\left[\begin{array}{cc|c} 1 & 1 & a \\ 0 & -2 & b - a \end{array}\right]$, de donde se puede ver que existe una solución (única). [Más operaciones con renglón producen $x = (a + b)/2$, $y = (a - b)/2$.] Por tanto, $\mathbb{R}^2 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$.

11. Es necesario demostrar que la ecuación vectorial $x\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$ tiene una solución para todos los valores de a, b y c. Esta ecuación vectorial es equivalente al sistema lineal cuya matriz aumentada es

$\left[\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 0 & 1 & 1 & b \\ 1 & 0 & 1 & c \end{array}\right]$. La reducción por renglón produce

$\left[\begin{array}{ccc|c} 1 & 1 & 0 & a \\ 0 & 1 & 1 & b \\ 0 & 0 & 2 & b + c - a \end{array}\right]$, de donde puede ver que existe una solución (única). [Más operaciones con renglón producen $x = (a - b + c)/2$, $y = (a + b - c)/2$, $z = (-a + b + c)/2$.]

Por tanto, $\mathbb{R}^3 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right)$.

13. **(a)** La recta a través del origen con vector director $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$
(b) La recta con ecuación general $2x + y = 0$

15. **(a)** El plano que pasa por el origen con vectores directores $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$
(b) El plano con ecuación general $2x - y + 4z = 0$

17. La sustitución produce el sistema lineal

$$\begin{aligned} a \qquad + 3c &= 0 \\ -a + b - 3c &= 0 \end{aligned}$$

cuya solución es $t\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$. Se tiene que existe un número infinito de soluciones, siendo acaso la más simple $a = -3$, $b = 0$, $c = 1$.

19. $\mathbf{u} = \mathbf{u} + 0(\mathbf{u} + \mathbf{v}) + 0(\mathbf{u} + \mathbf{v} + \mathbf{w})$
$\mathbf{v} = (-1)\mathbf{u} + (\mathbf{u} + \mathbf{v}) + 0(\mathbf{u} + \mathbf{v} + \mathbf{w})$
$\mathbf{w} = 0\mathbf{u} + (-1)(\mathbf{u} + \mathbf{v}) + (\mathbf{u} + \mathbf{v} + \mathbf{w})$

21. **(c)** Debe demostrar que gen$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ = gen$(\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3)$. Se sabe que gen$(\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3) \subseteq \mathbb{R}^3$ = gen$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$. Del ejercicio 19, $\mathbf{e}_1$, $\mathbf{e}_2$ y $\mathbf{e}_3$ todos pertenecen a gen$(\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3)$. Por tanto, por el ejercicio 21(b), gen$(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ = gen$(\mathbf{e}_1, \mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3)$.

23. Linealmente independiente

25. Linealmente dependiente, $-\begin{bmatrix}0\\1\\2\end{bmatrix} + \begin{bmatrix}2\\1\\3\end{bmatrix} = \begin{bmatrix}2\\0\\1\end{bmatrix}$

27. Linealmente dependiente, puesto que el conjunto contiene el vector cero

29. Linealmente independiente

31. Linealmente dependiente, $\begin{bmatrix}3\\-1\\1\\-1\end{bmatrix} + \begin{bmatrix}-1\\3\\1\\-1\end{bmatrix} + \begin{bmatrix}-1\\-1\\1\\3\end{bmatrix} = \begin{bmatrix}1\\1\\3\\1\end{bmatrix}$

43. **(a)** Sí **(b)** No

Ejercicios 2.4

1. $x_1 = 160, x_2 = 120, x_3 = 160$

3. dos pequeños, tres medianos, cuatro grandes

5. 65 bolsas de mezcla de la casa, 30 bolsas de mezcla especial, 45 bolsas de mezcla gourmet

7. $4FeS_2 + 11O_2 \longrightarrow 2Fe_2O_3 + 8SO_2$

9. $2C_4H_{10} + 13O_2 \longrightarrow 8CO_2 + 10H_2O$

11. $2C_5H_{11}OH + 15O_2 \longrightarrow 12H_2O + 10CO_2$

13. $Na_2CO_3 + 4C + N_2 \longrightarrow 2NaCN + 3CO$

15. **(a)** $f_1 = 30 - t$
$f_2 = -10 + t$
$f_3 = t$

(b) $f_1 = 15, f_3 = 15$

(c) $0 \leq f_1 \leq 20$
$0 \leq f_2 \leq 20$
$10 \leq f_3 \leq 30$

(d) Flujo negativo significaría que el agua fluye hacia atrás, contra la dirección de la flecha.

17. **(a)** $f_1 = -200 + s + t$
$f_2 = 300 - s - t$
$f_3 = s$
$f_4 = 150 - t$
$f_5 = t$

(b) $200 \leq f_3 \leq 300$

(c) Si $f_3 = s = 0$, entonces $f_5 = t \geq 200$ (de la ecuación f_1), pero $f_5 = t \leq 150$ (de la ecuación f_4). Esto es una contradicción.

(d) $50 \leq f_3 \leq 300$

19. $I_1 = 3$ amps, $I_2 = 5$ amps, $I_3 = 2$ amps

21. **(a)** $I = 10$ amps, $I_1 = I_5 = 6$ amps, $I_2 = I_4 = 4$ amps, $I_3 = 2$ amps

(b) $R_{eff} = \frac{7}{5}$ ohms

(c) Sí; lo cambia a 4 ohms.

23. Agrícola: Manufacturera = 2 : 3

25. El pintor cobra \$39/h, el plomero \$42/h, el electricista \$54/h.

27. **(a)** El carbón debe producir \$100 millones y el acero \$160 millones.

(b) El carbón debe reducir la producción en ≈ \$4.2 millones y el acero debe aumentar la producción en ≈ \$5.7 millones.

29. **(a)** Sí; presione los interruptores 1, 2 y 3 o los interruptores 3, 4 y 5.

(b) No

31. Los estados que pueden obtenerse se representan mediante estos vectores

$$\begin{bmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{bmatrix}$$

en $\mathbb{Z}_2^5$ para el cual $x_1 + x_2 + x_4 + x_5 = 0$.
(Existen 16 de tales posibilidades.)

33. Si 0 = apagada, 1 = azul claro y 2 = azul oscuro, entonces el sistema lineal que surge tiene matriz aumentada

$$\left[\begin{array}{ccccc|c} 1 & 1 & 0 & 0 & 0 & 2 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \end{array}\right]$$

que se reduce sobre $\mathbb{Z}_3$ a

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Esto produce las soluciones

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 2 \\ 1 \\ 0 \\ 2 \\ 1 \end{bmatrix}$$

donde t está en $\mathbb{Z}_3$. Por tanto, existen exactamente tres soluciones:

$$\begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 2 \end{bmatrix}$$

donde cada entrada indica el número de veces que debe oprimir el interruptor correspondiente.

35. **(a)** Oprima los cuadrados 3 y 7.
(b) La matriz de coeficientes A de 9×9 es equivalente por renglones a $\mathbb{Z}_2$, de modo que para cualquier $\mathbf{b}$ en $\mathbb{Z}_2^9$, $A\mathbf{x} = \mathbf{b}$ tiene una solución única.

37. Grace tiene 15 y Hans 5.

39. 1200 y 600 yardas cuadradas.

41. **(a)** $a = 4 - d, b = 5 - d, c = -2 + d, d$ es arbitrario
(b) No hay solución

43. **(a)** No hay solución
(b) $[a, b, c, d, e, f] = [4, 5, 6, -3, -1, 0] + f[-1, -1, -1, 1, 1, 1]$

45. **(a)** $y = x^2 - 2x + 1$ **(b)** $y = x^2 + 6x + 10$

47. $A = 1, B = 2$

49. $A = -\frac{1}{5}, B = \frac{1}{3}, C = 0, D = -\frac{2}{15}, E = -\frac{1}{5}$

51. $a = \frac{1}{2}, b = \frac{1}{2}, c = 0$

Ejercicios 2.5

1.

n	0	1	2	3	4	5
x_1	0	0.8571	0.9714	0.9959	0.9991	0.9998
x_2	0	0.8000	0.9714	0.9943	0.9992	0.9998

Solución exacta: $x_1 = 1, x_2 = 1$

3.

n	0	1	2	3	4	5	6
x_1	0	0.2222	0.2539	0.2610	0.2620	0.2622	0.2623
x_2	0	0.2857	0.3492	0.3582	0.3603	0.3606	0.3606

Solución exacta (a cuatro lugares decimales): $x_1 = 0.2623, x_2 = 0.3606$

5.

n	0	1	2	3	4	5	6	7	8
x_1	0	0.3333	0.2500	0.3055	0.2916	0.3009	0.2986	0.3001	0.2997
x_2	0	0.2500	0.0834	0.1250	0.0972	0.1042	0.0996	0.1008	0.1000
x_3	0	0.3333	0.2500	0.3055	0.2916	0.3009	0.2986	0.3001	0.2997

Solución exacta: $x_1 = 0.3, x_2 = 0.1, x_3 = 0.3$

7.

n	0	1	2	3	4
x_1	0	0.8571	0.9959	0.9998	1.0000
x_2	0	0.9714	0.9992	1.0000	1.0000

Después de tres iteraciones, el método de Gauss-Seidel está dentro de 0.001 de la solución exacta. El método de Jacobi requiere cuatro iteraciones para alcanzar la misma precisión.

9.

n	0	1	2	3	4
x_1	0	0.2222	0.2610	0.2622	0.2623
x_2	0	0.3492	0.3603	0.3606	0.3606

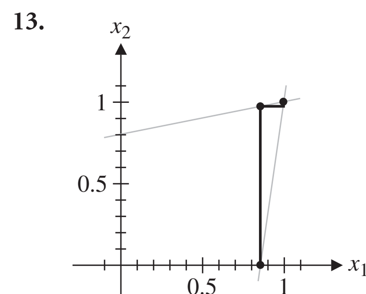
Después de tres iteraciones, el método de Gauss-Seidel está dentro de 0.001 de la solución exacta. El método de Jacobi requiere cuatro iteraciones para alcanzar la misma precisión.

11.

n	0	1	2	3	4	5	6
x_1	0	0.3333	0.2777	0.2962	0.2993	0.2998	0.3000
x_2	0	0.1667	0.1112	0.1020	0.1004	0.1000	0.1000
x_3	0	0.2777	0.2962	0.2993	0.2998	0.3000	0.3000

Después de cuatro iteraciones, el método de Gauss-Seidel está dentro de 0.001 de la solución exacta. El método de Jacobi requiere siete iteraciones para alcanzar la misma precisión.

13.



15.

n	0	1	2	3	4
x_1	0	3	-5	19	-53
x_2	0	-4	8	-28	80

Si las ecuaciones se intercambian y se aplica el método de Gauss-Seidel al sistema equivalente

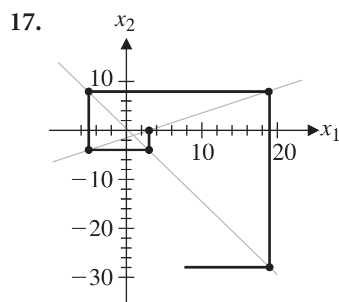
$$3x_1 + 2x_2 = 1$$
$$x_1 - 2x_2 = 3$$

se obtiene

n	0	1	2	3	4	5	6	7	8
x_1	0	0.3333	1.2222	0.9260	1.0247	0.9918	1.0027	0.9991	1.0003
x_2	0	-1.3333	-0.8889	-1.0370	-0.9876	-1.0041	-0.9986	-1.0004	-0.9998

Después de siete iteraciones, el proceso converge hasta dentro de 0.001 de la solución exacta $x_1 = 1, x_2 = -1$.

17.



19.

n	0	1	2	3	4	5	6
x_1	0	-1.6	14.97	8.550	10.740	9.839	10.120
x_2	0	25.9	11.408	14.051	11.615	11.718	11.249
x_3	0	-10.35	-9.311	-11.200	-11.322	-11.721	-11.816

n	7	8	9	10	11	12
x_1	9.989	10.022	10.002	10.005	10.001	10.001
x_2	11.187	11.082	11.052	11.026	11.015	11.008
x_3	-11.912	-11.948	-11.973	-11.985	-11.992	-11.996

Después de 12 iteraciones, el método de Gauss-Seidel converge dentro de 0.01 de la solución exacta $x_1 = 10, x_2 = 11, x_3 = -12$.

21.

n	13	14	15	16
x_1	10.0004	10.0003	10.0001	10.0001
x_2	11.0043	11.0023	11.0014	11.0007
x_3	−11.9976	−11.9986	−11.9993	−11.9996

23. El método de Gauss-Seidel produce

n	0	1	2	3	4	5	6	7	8	9
x_1	0	0	12.5	21.875	24.219	24.805	24.951	24.988	24.997	24.999
x_2	0	0	18.75	21.438	24.609	24.902	24.976	24.994	24.998	24.999
x_3	0	50	68.75	73.438	74.609	74.902	74.976	74.994	74.998	74.999
x_4	0	62.5	71.875	74.219	74.805	74.951	74.988	74.997	74.999	75.000

La solución exacta es $x_1 = 25, x_2 = 25, x_3 = 75, x_4 = 75$.

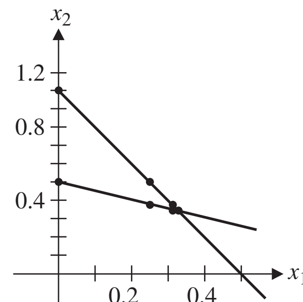
25. El método de Gauss-Seidel produce las siguientes iteraciones:

n	0	1	2	3	4	5	6
t_1	0	20	21.25	22.8125	23.3301	23.6596	23.7732
t_2	0	5	11.25	13.3203	14.6386	15.0926	15.2732
t_3	0	21.25	24.6094	26.9873	27.7303	27.9626	28.0352
t_4	0	2.5	5.8594	8.2373	8.9804	9.2126	9.2852
t_5	0	7.1875	14.6289	16.2829	16.7578	16.9036	16.9491
t_6	0	23.0469	24.9072	25.3207	25.4394	25.4759	25.4873

n	7	8	9	10	11	12
t_1	23.8093	23.8206	23.8242	23.8252	23.8256	23.8257
t_2	15.2824	15.2966	15.3010	15.3024	15.3029	15.3029
t_3	28.0579	28.0650	28.0671	28.0678	28.0681	28.0681
t_4	9.3079	9.3150	9.3172	9.3178	9.3181	9.3181
t_5	16.9633	16.9677	16.9690	16.9695	16.9696	16.9696
t_6	25.4908	25.4919	25.4922	25.4924	25.4924	25.4924

27. (a)

n	0	1	2	3	4	5	6
x_1	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{5}{16}$	$\frac{5}{16}$	$\frac{21}{64}$
x_2	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{11}{32}$	$\frac{11}{32}$



(b) $2x_1 + x_2 = 1$
$x_1 + 2x_2 = 1$

(c)

n	0	1	2	3	4	5	6	7
x_1	0	0	0.25	0.3125	0.3281	0.3320	0.3330	0.3332
x_2	1	0.5	0.375	0.3438	0.3360	0.3340	0.3335	0.3334

[Las columnas 1, 2 y 3 de esta tabla son las columnas *impares* 1, 3 y 5 de la tabla en el inciso (a). Las iteraciones convergen en $x_1 = x_2 = 0.3333$.

(d) $x_1 = x_2 = \frac{1}{3}$

Preguntas de repaso

1. (a) F (c) F (e) V (g) V (i) F

3. $\begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix}$ **5.** $\begin{bmatrix} 6 \\ 2 \end{bmatrix}$ **7.** $k = -1$ **9.** $(0, 3, 1)$

11. $x - 2y + z = 0$ **13.** (a) Si **15.** 1 o 2

17. Si $c_1(\mathbf{u} + \mathbf{v}) + c_2(\mathbf{u} - \mathbf{v}) = \mathbf{0}$, entonces $(c_1 + c_2)\mathbf{u} + (c_1 - c_2)\mathbf{v} = \mathbf{0}$. La independencia lineal de $\mathbf{u}$ y $\mathbf{v}$ implica $c_1 + c_2 = 0$ y $c_1 - c_2 = 0$. Al resolver este sistema, se obtiene $c_1 = c_2 = 0$. Por tanto, $\mathbf{u} + \mathbf{v}$ y $\mathbf{u} - \mathbf{v}$ son linealmente independientes.

19. Sus rangos deben ser iguales.

Capítulo 3

Ejercicios 3.1

1. $\begin{bmatrix} 3 & -6 \\ -5 & 7 \end{bmatrix}$ **3.** No es posible

5. $\begin{bmatrix} 12 & -6 & 3 \\ -4 & 12 & 14 \end{bmatrix}$ **7.** $\begin{bmatrix} 3 & 3 \\ 19 & 27 \end{bmatrix}$

9. $[10]$ **11.** $\begin{bmatrix} -4 & -2 \\ 8 & 4 \end{bmatrix}$

13. $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ **15.** $\begin{bmatrix} 27 & 0 \\ -49 & 125 \end{bmatrix}$

17. $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

19. $B = \begin{bmatrix} 1.50 & 1.00 & 2.00 \\ 1.75 & 1.50 & 1.00 \end{bmatrix}$, $BA = \begin{bmatrix} 650.00 & 462.50 \\ 675.00 & 406.25 \end{bmatrix}$

La columna i corresponde al almacén i, el renglón 1 contiene los costos de embarque por camión y el renglón 2 contiene los costos de embarcar por tren.

21. $\begin{bmatrix} 1 & -2 & 3 \\ 2 & 1 & -5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 4 \end{bmatrix}$

23. $AB = [2\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_3 \quad 3\mathbf{a}_1 - \mathbf{a}_2 + 6\mathbf{a}_3 \quad \mathbf{a}_2 + 4\mathbf{a}_3]$ (donde $\mathbf{a}_i$ es la i-ésima columna de A)

25. $\begin{bmatrix} 2 & 3 & 0 \\ -6 & -9 & 0 \\ 4 & 6 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & -12 & -8 \\ -1 & 6 & 4 \\ 1 & -6 & -4 \end{bmatrix}$

27. $BA = \begin{bmatrix} 2\mathbf{A}_1 + 3\mathbf{A}_2 \\ \mathbf{A}_1 - \mathbf{A}_2 + \mathbf{A}_3 \\ -\mathbf{A}_1 + 6\mathbf{A}_2 + 4\mathbf{A}_3 \end{bmatrix}$ (donde $\mathbf{A}_i$ es el i-ésimo renglón de A)

29. Si $\mathbf{b}_i$ es la i-ésima columna de B, entonces $A\mathbf{b}_i$ es la i-ésima columna de AB. Si las columnas de B son linealmente dependientes, entonces existen escalares $c_1, \ldots, c_n$ (no todos cero) tales que $c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n = \mathbf{0}$. Pero entonces $c_1(A\mathbf{b}_1) + \cdots + c_n(A\mathbf{b}_n) = A(c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n) = A\mathbf{0} = \mathbf{0}$, de modo que las columnas de AB son linealmente dependientes.

31. $\left[\begin{array}{cc|c} 3 & 2 & 0 \\ -1 & 1 & 0 \\ \hline 0 & 0 & 5 \end{array}\right]$ **33.** $\left[\begin{array}{cc|cc} 1 & 2 & 2 & 0 \\ 3 & 4 & 5 & 3 \\ \hline 1 & 0 & 1 & 2 \\ 0 & 1 & 0 & -1 \end{array}\right]$

35. (a) $A^2 = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$, $A^3 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$,

$A^4 = \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$, $A^5 = \begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix}$, $A^6 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,

$A^7 = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$

(b) $A^{2001} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

37. $A^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$

39. (a) $\begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{bmatrix}$ (c) $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 2 & 4 & 8 & 16 \\ 3 & 9 & 27 & 81 \end{bmatrix}$

Ejercicios 3.2

1. $X = \begin{bmatrix} 5 & 4 \\ 3 & 5 \end{bmatrix}$ **3.** $X = \begin{bmatrix} -\frac{2}{3} & \frac{4}{3} \\ \frac{10}{3} & 4 \end{bmatrix}$

5. $B = 2A_1 + A_2$ **7.** No es posible

9. $\text{gen}(A_1, A_2) = \left\{\begin{bmatrix} c_1 & 2c_1 + c_2 \\ -c_1 + 2c_2 & c_1 + c_2 \end{bmatrix}\right\} =$
$\left\{\begin{bmatrix} w & x \\ 2x - 5w & x - w \end{bmatrix}\right\}$

11. $\text{gen}(A_1, A_2, A_3) =$
$\left\{\begin{bmatrix} c_1 - c_2 + c_3 & 2c_2 + c_3 & -c_1 + c_3 \\ 0 & c_1 + c_2 & 0 \end{bmatrix}\right\} =$
$\left\{\begin{bmatrix} -3b + 4c + 5e & b & c \\ 0 & e & 0 \end{bmatrix}\right\}$

13. Linealmente independiente

15. Linealmente independiente

23. $a = d, c = 0$ **25.** $3b = 2c, a = d - c$

27. $a = d, b = c = 0$

29. Sean $A = [a_{ij}]$ y $B = [b_{ij}]$ matrices triangulares superiores de $n \times n$ y sea $i > j$. Entonces, por la definición de matriz triangular superior,
$a_{i1} = a_{i2} = \cdots = a_{i,\,i-1} = 0$ y
$b_{ij} = b_{i+1,\,j} = \cdots = b_{nj} = 0$
Ahora sea $C = AB$. Entonces

$$\begin{aligned} c_{ij} = {} & a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots a_{i,\,i-1}b_{i-1,\,j} + a_{ii}b_{ij} \\ & + a_{i,\,i+1}b_{i+1,\,j} + \cdots + a_{in}b_{nj} \\ = {} & 0 \cdot b_{1j} + 0 \cdot b_{2j} + \cdots 0 \cdot b_{i-1,\,j} + a_{ii} \cdot 0 \\ & + a_{i,\,i+1} \cdot 0 + \cdots + a_{in} \cdot 0 = 0 \end{aligned}$$

de donde se concluye que C es triangular superior.

35. (a) A, B simétrica $\Rightarrow (A + B)^T = A^T + B^T = A + B \Rightarrow A + B$ es simétrica

37. Las matrices (b) y (c) son antisimétricas.

41. A o B (o ambas) deben ser la matriz cero.

43. (b) $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 3 & 5 \\ 3 & 5 & 7 \\ 5 & 7 & 9 \end{bmatrix} + \begin{bmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{bmatrix}$

47. *Sugerencia:* use la traza.

Ejercicios 3.3

1. $\begin{bmatrix} 2 & -7 \\ -1 & 4 \end{bmatrix}$ **3.** No es invertible

5. No es invertible **7.** $\begin{bmatrix} -1.6 & -2.8 \\ 0.\overline{3} & 1 \end{bmatrix}$

9. $\begin{bmatrix} a/(a^2 + b^2) & b/(a^2 + b^2) \\ -b/(a^2 + b^2) & a/(a^2 + b^2) \end{bmatrix}$ **11.** $\begin{bmatrix} -5 \\ 9 \end{bmatrix}$

13. (a) $\mathbf{x}_1 = \begin{bmatrix} 4 \\ -\frac{1}{2} \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} -5 \\ 2 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$

(c) El método en el inciso (b) usa menos multiplicaciones

17. (b) $(AB)^{-1} = A^{-1}B^{-1}$ si y sólo si $AB = BA$

21. $X = A^{-1}(BA)^2B^{-1}$

23. $X = (AB)^{-1}BA + A$

25. $E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ **27.** $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

29. $E = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ **31.** $\begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{bmatrix}$

33. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ **35.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ **37.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/c & 0 \\ 0 & 0 & 1 \end{bmatrix}$

39. $A = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}, A^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{2} \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$

43. (a) Si A es invertible, entonces $BA = CA \Rightarrow (BA)A^{-1} = (CA)A^{-1} \Rightarrow B(AA^{-1}) = C(AA^{-1}) \Rightarrow BI = CI \Rightarrow B = C$.

45. *Sugerencia:* reescriba $A^2 - 2A + I = O$ como $A(2I - A) = I$.

47. Si AB es invertible, entonces existe una matriz X tal que $(AB)X = I$. Pero entonces $A(BX) = I$ también, de modo que A es invertible (con inverso BX).

49. $\begin{bmatrix} \frac{1}{10} & \frac{2}{5} \\ \frac{3}{10} & \frac{1}{5} \end{bmatrix}$ **51.** $\begin{bmatrix} 1/(a^2 + 1) & -a/(a^2 + 1) \\ a/(a^2 + 1) & 1/(a^2 + 1) \end{bmatrix}$

53. No es invertible

55. $\begin{bmatrix} 1/a & 0 & 0 \\ -1/a^2 & 1/a & 0 \\ 1/a^3 & -1/a^2 & 1/a \end{bmatrix}, a \neq 0$

57. $\begin{bmatrix} -11 & -2 & 5 & -4 \\ 4 & 1 & -2 & 2 \\ 5 & 1 & -2 & 2 \\ 9 & 2 & -4 & 3 \end{bmatrix}$

59. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -a/d & -b/d & -c/d & 1/d \end{bmatrix}, d \neq 0$

61. No es invertible **63.** $\begin{bmatrix} 4 & 6 & 4 \\ 5 & 3 & 2 \\ 0 & 6 & 5 \end{bmatrix}$

69. $\left[\begin{array}{cc|cc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline -2 & -3 & 1 & 0 \\ -1 & -2 & 0 & 1 \end{array}\right]$ **71.** $\left[\begin{array}{cc|cc} -1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{array}\right]$

Ejercicios 3.4

1. $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ **3.** $\begin{bmatrix} -3/2 \\ -2 \\ -1 \end{bmatrix}$ **5.** $\begin{bmatrix} -7 \\ -15 \\ -2 \\ 2 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 0 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 5 \end{bmatrix}$

9. $\begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 8 & 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 3 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ -1 & 0 & -2 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 & -1 \\ 0 & 2 & -3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 & -2 \\ 0 & 3 & 3 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix}$

15. $L^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, U^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{12} \\ 0 & \frac{1}{6} \end{bmatrix}, A^{-1} = \begin{bmatrix} -5/12 & 1/12 \\ 1/6 & 1/6 \end{bmatrix}$

19. $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

21. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

23. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 5 & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & -16 \end{bmatrix}$

25. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 1 & 1 & 2 \\ 0 & -1 & 1 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}$

27. $\begin{bmatrix} 4 \\ -1 \\ -2 \end{bmatrix}$ **31.** $\begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} -2 & 0 \\ 0 & 6 \end{bmatrix}\begin{bmatrix} 1 & -\frac{1}{2} \\ 0 & 1 \end{bmatrix}$

Ejercicios 3.5

1. Subespacio **3.** Subespacio

5. Subespacio **7.** No es subespacio

11. $\mathbf{b}$ está en col(A), $\mathbf{w}$ no está en renglón(A).

15. No

17. $\{[1 \quad 0 \quad -1], [0 \quad 1 \quad 2]\}$ es una base para renglón(A); $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ es una base par col(A); $\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right\}$ es una base para nulo(A).

19. $\{[1 \quad 0 \quad 1 \quad 0], [0 \quad 1 \quad -1 \quad 0], [0 \quad 0 \quad 0 \quad 1]\}$ es una base para renglón(A); $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}\right\}$ es una base para col(A); $\left\{\begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}\right\}$ es una base para nulo(A).

21. $\{[1 \quad 0 \quad -1], [1 \quad 1 \quad 1]\}$ es una base para renglón(A); $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ es una base para col(A)

23. $\{[1 \quad 1 \quad 0 \quad 1], [0 \quad 1 \quad -1 \quad 1], [0 \quad 1 \quad -1 \quad -1]\}$ es una base para renglón(A); $\left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$ es una base para col(A)

25. Tanto $\{[1 \quad 0 \quad -1], [0 \quad 1 \quad 2]\}$ como $\{[1 \quad 0 \quad -1], [1 \quad 1 \quad 1]\}$ son conjuntos generadores linealmente independientes para renglón$(A) = \{[a \quad b \quad -a + 2b]\}$. Tanto $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ como $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ son conjuntos generadores linealmente independientes para col$(A) = \mathbb{R}^2$.

27. $\left\{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$

29. $\{[1 \quad 0 \quad 0], [0 \quad 1 \quad 0], [0 \quad 0 \quad 1]\}$

31. $\{[2 \quad -3 \quad 1], [1 \quad -1 \quad 0], [4 \quad -4 \quad 1]\}$

35. rank$(A) = 2$, nulidad$(A) = 1$

37. rank$(A) = 3$, nulidad$(A) = 1$

39. Si A es de 3×5 entonces rank$(A) \leq 3$, de modo que no puede haber más de tres columnas linealmente independientes.

41. nulidad$(A) = 2, 3, 4$ o 5

43. Si $a = -1$, entonces rank$(A) = 1$; si $a = 2$, entonces rank$(A) = 2$; para $a \neq -1, 2$, rank$(A) = 3$.

45. Sí **47.** Sí **49.** No

51. $\mathbf{w}$ está en gen$(\mathcal{B})$ si y sólo si el sistema lineal con matriz aumentada $[\mathcal{B} \mid \mathbf{w}]$ es consistente, lo que es verdadero en este caso, pues

$$[\mathcal{B} \mid \mathbf{w}] = \left[\begin{array}{rr|r} 1 & 1 & 1 \\ 2 & 0 & 6 \\ 0 & -1 & 2 \end{array}\right] \longrightarrow \left[\begin{array}{rr|r} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{array}\right]$$

A partir de esta forma escalonada reducida por renglón, también es claro que $[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

53. rank$(A) = 2$, nulidad$(A) = 1$

55. rank$(A) = 3$, nulidad$(A) = 1$

57. Sean $\mathbf{A}_1, \ldots, \mathbf{A}_m$ los vectores renglón de A, de modo que renglón $(A) =$ gen$(\mathbf{A}_1, \ldots, \mathbf{A}_m)$. Si $\mathbf{x}$ está en nulo(A), entonces, dado que $A\mathbf{x} = 0$, también se tiene $\mathbf{A}_i \cdot \mathbf{x} = 0$ para $i = 1 \ldots, m$, por la definición renglón-columna de multiplicación matricial. Si $\mathbf{r}$ está en renglón(A), entonces $\mathbf{r}$ es de la forma $\mathbf{r} = c_1\mathbf{A}_1 + \cdots + c_m\mathbf{A}_m$. Por tanto,

$$\begin{aligned} \mathbf{r} \cdot \mathbf{x} &= (c_1\mathbf{A}_1 + \cdots + c_m\mathbf{A}_m) \cdot \mathbf{x} \\ &= c_1(\mathbf{A}_1 \cdot \mathbf{x}) + \cdots + c_m(\mathbf{A}_m \cdot \mathbf{x}) = 0 \end{aligned}$$

59. (a) Si un conjunto de columnas de AB es linealmente independiente, entonces las correspondientes columnas de B son linealmente independientes (por un argumento similar al necesario para probar el ejercicio 29 de la sección 3.1). Se tiene que el *máximo* número k de columnas linealmente independientes de AB [esto es, $k =$ rank(AB)] no es más que el *máximo* número r de columnas linealmente independientes de B [esto es, $r =$ rank(B)]. En otras palabras, rank$(AB) \leq$ rank(B).

61. (a) Del ejercicio 59(a), rank(UA)$\leq$ rank(A) y rank$(A) =$ rank$((U^{-1}U)A) =$ rank$(U^{-1}(UA)) \leq$ rank(UA). Por tanto, rank$(UA) =$ rank(A).

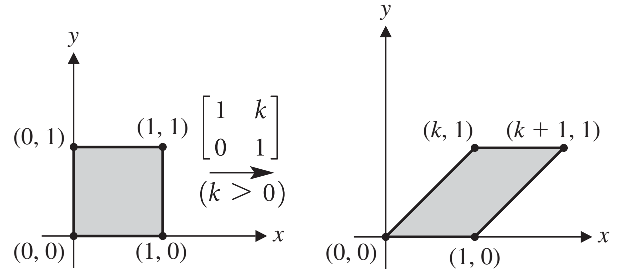
Ejercicios 3.6

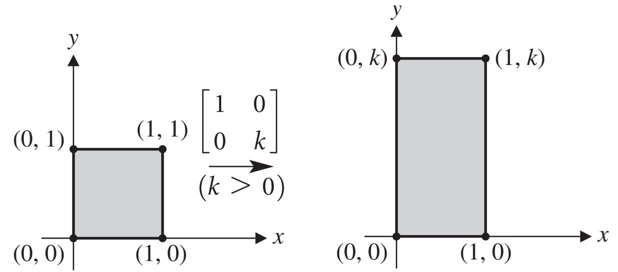
1. $T(\mathbf{u}) = \begin{bmatrix} 0 \\ 11 \end{bmatrix}$, $T(\mathbf{v}) = \begin{bmatrix} 8 \\ 1 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ **13.** $\begin{bmatrix} 1 & -1 & 1 \\ 2 & 1 & -3 \end{bmatrix}$

15. $[F] = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ **17.** $[D] = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

19. $\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$ estira o contrae en la dirección x (combinado con una reflexión en el eje y si $k < 0$); $\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$ estira o contrae en la dirección y (combinado con una reflexión en el eje x si $k < 0$); $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ es una reflexión en la recta $y = x$; $\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ es un *corte* en la dirección x; $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$ es un *corte* en la dirección y. Por ejemplo,





21. $\begin{bmatrix} \sqrt{3}/2 & 1/2 \\ -1/2 & \sqrt{3}/2 \end{bmatrix}$ **23.** $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$

25. $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ **27.** $\begin{bmatrix} -\frac{3}{5} & \frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix}$

31. $[S \circ T] = \begin{bmatrix} -8 & 5 \\ 4 & 1 \end{bmatrix}$

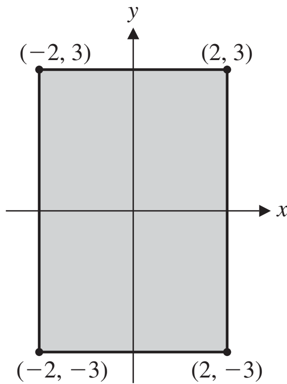
33. $[S \circ T] = \begin{bmatrix} 0 & 6 & -6 \\ 1 & -2 & 2 \end{bmatrix}$

35. $[S \circ T] = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$

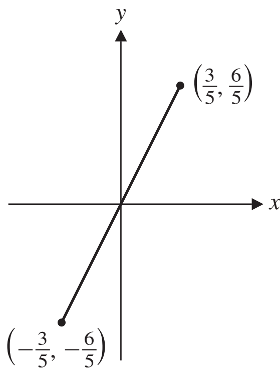
37. $\begin{bmatrix} -\sqrt{3}/2 & 1/2 \\ 1/2 & \sqrt{3}/2 \end{bmatrix}$ **39.** $\begin{bmatrix} -\sqrt{3}/2 & -1/2 \\ 1/2 & -\sqrt{3}/2 \end{bmatrix}$

45. En forma vectorial, sean las rectas paralelas dadas por $\mathbf{x} = \mathbf{p} + t\mathbf{d}$ y $\mathbf{x}' = \mathbf{p}' + t\mathbf{d}$. Sus imágenes son $T(\mathbf{x}) = T(\mathbf{p} + t\mathbf{d}) = T(\mathbf{p}) + tT(\mathbf{d})$ y $T(\mathbf{x}') = T(\mathbf{p}' + t\mathbf{d}) = T(\mathbf{p}') + tT(\mathbf{d})$. Suponga que $T(\mathbf{d}) \neq \mathbf{0}$. Si $T(\mathbf{p}') - T(\mathbf{p})$ es paralela a $T(\mathbf{d})$, entonces las imágenes representan la misma recta; de otro modo, las imágenes representan distintas rectas paralelas. Por otra parte, si $T(\mathbf{d}) = \mathbf{0}$, entonces las imágenes representan dos puntos distintos si $T(\mathbf{p}') \neq T(\mathbf{p})$ y un punto individual de otra manera.

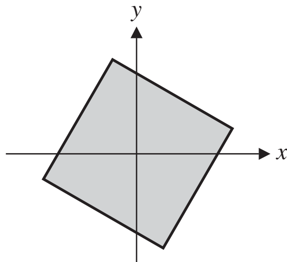
47.



49.



51.



Ejercicios 3.7

1. $\mathbf{x}_1 = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0.38 \\ 0.62 \end{bmatrix}$

3. 64%

5. $\mathbf{x}_1 = \begin{bmatrix} 150 \\ 120 \\ 120 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 155 \\ 120 \\ 115 \end{bmatrix}$

7. $\frac{5}{18}$

9. **(a)** $P = \begin{bmatrix} 0.662 & 0.250 \\ 0.338 & 0.750 \end{bmatrix}$ **(b)** 0.353

(c) 42.5% húmedo, 57.5% seco

11. **(a)** $P = \begin{bmatrix} 0.08 & 0.09 & 0.11 \\ 0.07 & 0.11 & 0.05 \\ 0.85 & 0.80 & 0.84 \end{bmatrix}$

(b) 0.08, 0.1062, 0.1057, 0.1057, 0.1057

(c) 10.6% bueno, 5.5% adecuado, 83.9% pobre

13. Las entradas del vector $\mathbf{j}P$ son justo las sumas de columna de la matriz P. De modo que P es estocástica si y sólo si $\mathbf{j}P = \mathbf{j}$.

15. 4

17. 9.375

19. Sí, $\mathbf{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

21. No

23. No

25. Sí, $\mathbf{x} = \begin{bmatrix} 10 \\ 27 \\ 35 \end{bmatrix}$

27. Productiva

29. No productiva

31. $\mathbf{x} = \begin{bmatrix} 10 \\ 16 \end{bmatrix}$

33. Sí, $\mathbf{x} = \begin{bmatrix} 10 \\ 6 \\ 8 \end{bmatrix}$

37. $\mathbf{x}_1 = \begin{bmatrix} 500 \\ 70 \\ 50 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 720 \\ 350 \\ 35 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 1175 \\ 504 \\ 175 \end{bmatrix}$

39. **(a)** Para L_1 se tiene $\mathbf{x}_1 = \begin{bmatrix} 50 \\ 8 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 40 \\ 40 \end{bmatrix},$

$\mathbf{x}_3 = \begin{bmatrix} 200 \\ 32 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 160 \\ 160 \end{bmatrix}, \mathbf{x}_5 = \begin{bmatrix} 800 \\ 128 \end{bmatrix}, \mathbf{x}_6 = \begin{bmatrix} 640 \\ 640 \end{bmatrix},$

$\mathbf{x}_7 = \begin{bmatrix} 3200 \\ 512 \end{bmatrix}, \mathbf{x}_8 = \begin{bmatrix} 2560 \\ 2560 \end{bmatrix}, \mathbf{x}_9 = \begin{bmatrix} 12800 \\ 2048 \end{bmatrix},$

$\mathbf{x}_{10} = \begin{bmatrix} 10240 \\ 10240 \end{bmatrix}.$

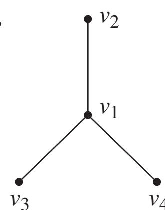
(b) La primera población oscila entre dos estados, mientras que la segunda tiende a un estado estacionario.

41. La población oscila a lo largo de un ciclo de tres estados (para la población relativa): si $0.1 < s \le 1$, la población real crece; si $s = 0.1$, la población real pasa por un ciclo de longitud 3; y si $0 \le s < 0.1$, la población real declina (y a la larga morirá).

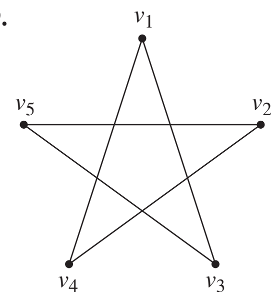
43. $A = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$

45. $A = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$

47.

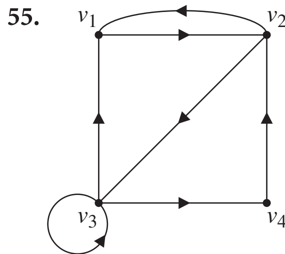


49.

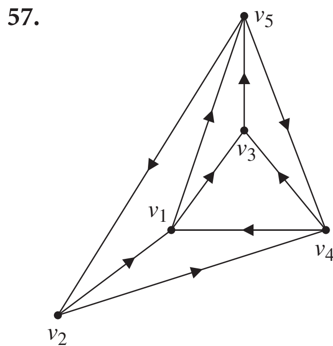


51. $A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$ **53.** $A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$

55.



57.

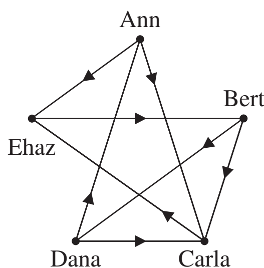


59. 2 **61.** 3 **63.** 0 **65.** 3

67. **(a)** El vértice i no es adyacente a algún otro vértice.

69. Si sólo se usan victorias directas, P_2 está en primer lugar; P_3, P_4 y P_6 empatan en segundo lugar; y P_1 y P_5 empatan en tercer lugar. Si se combinan victorias directas e indirectas, los jugadores se clasifican del modo siguiente: P_2 en primer lugar, seguido por P_6, P_4, P_3, P_5 y P_1.

71. **(a)**



$A = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}$

(b) dos pasos; todas las entradas fuera de la diagonal del segundo renglón de $A + A^2$ son distintas de cero.

(d) Si la gráfica tiene n vértices, compruebe que la entrada (i, j) de las potencias A^k para $k = 1, \ldots, n - 1$. El vértice i está conectado al vértice j mediante una trayectoria de longitud k si y sólo si $(A^k)_{ij} \neq 0$.

73. $(AA^T)_{ij}$ cuenta el número de vértices adyacentes a *ambos* vértices, el i y el j.

75. Bipartita **77.** Bipartita

79. Un solo error podría cambiar el vector código $\mathbf{c}_2 = [0, 1, 0, 1]$ en $\mathbf{c}' = [1, 1, 0, 1]$. Sin embargo, $\mathbf{c}'$ también podría obtenerse del vector código $\mathbf{c}_4 = [1, 1, 1, 1]$ vía un solo error, de modo que el error no puede corregirse.

81. $\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$ **83.** $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

85. $P\mathbf{c}' = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ es la segunda columna de P, de modo que el error está en el segundo componente de $\mathbf{c}'$. Por tanto, el vector de mensaje correcto (de los primeros cuatro componentes del vector corregido) es $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$.

87. **(a)** $P = [1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1 \;\; 1]$

(b) $G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$

(c) Las columnas de P no son distintas.

89. **(a)** $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$

(b) $P = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$; no es un código de corrección de error.

91. Un conjunto de candidatos para P y G es

$$P = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

y

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Preguntas de repaso

1. **(a)** V **(c)** F **(e)** V **(g)** V **(i)** V

3. Imposible

5. $\begin{bmatrix} \frac{17}{83} & -\frac{1}{83} \\ -\frac{1}{83} & \frac{5}{166} \end{bmatrix}$

7. $\begin{bmatrix} 1 & 3 \\ 3 & 9 \end{bmatrix} + \begin{bmatrix} 4 & 10 \\ 10 & 25 \end{bmatrix}$

9. $\begin{bmatrix} 0 & -9 \\ 2 & 4 \\ 1 & -6 \end{bmatrix}$

11. Porque $(I - A)(I + A + A^2) = I - A^3 = I - O = I$, $(I - A)^{-1} = I + A + A^2$.

13. Una base para renglón(A) es $\{[1, -2, 0, -1, 0], [0, 0, 1, 2, 0], [0, 0, 0, 0, 1]\}$; una base para col$(A)$ es

$$\left\{ \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \\ 6 \end{bmatrix} \right\}$$ (o la base estándar para $\mathbf{R}^3$); y una

base para nulo(A) es $\left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} \right\}$.

15. Una matriz invertible tiene un espacio nulo trivial (cero). Si A es invertible, entonces también lo es A^T, y en consecuencia tanto A como A^T tienen espacios nulos triviales. Si A no es invertible, entonces A y A^T no necesitan tener el mismo espacio nulo. Por ejemplo, considere

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}.$$

17. Puesto que A tiene n columnas linealmente independientes, rank$(A) = n$. Por tanto, rank$(A^TA) = n$ por el Teorema 3.28. Puesto que A^TA es de $n \times n$, esto implica que A^TA es invertible, por el teorema fundamental de las matrices invertibles. AA^T no necesita ser invertible.

Por ejemplo, considere $A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

19. $\begin{bmatrix} -1/5\sqrt{2} & -3/5\sqrt{2} \\ 2/5\sqrt{2} & 6/5\sqrt{2} \end{bmatrix}$

Capítulo 4

Ejercicios 4.1

1. $A\mathbf{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\mathbf{v}, \lambda = 3$

3. $A\mathbf{v} = \begin{bmatrix} -3 \\ 6 \end{bmatrix} = -3\mathbf{v}, \lambda = -3$

5. $A\mathbf{v} = \begin{bmatrix} 6 \\ -3 \\ 3 \end{bmatrix} = 3\mathbf{v}, \lambda = 3$

7. $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$ **9.** $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ **11.** $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$

13. $\lambda = 1, E_1 = \text{gen}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right); \lambda = -1, E_{-1} = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$

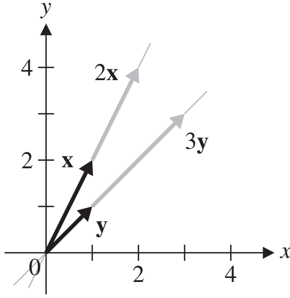
15. $\lambda = 0, E_0 = \text{gen}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right); \lambda = 1, E_1 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$

17. $\lambda = 2, E_2 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right); \lambda = 3, E_3 = \text{gen}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)$

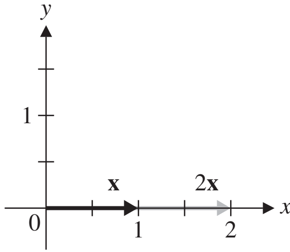
19. $\mathbf{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \lambda = 1; \mathbf{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \lambda = 2$

21. $\mathbf{v} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \lambda = 2; \mathbf{v} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \lambda = 0$

23. $\lambda = 2, E_2 = \text{gen}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right); \lambda = 3, E_3 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$



25. $\lambda = 2, E_2 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right)$



27. $\lambda = 1 + i, E_{1+i} = \text{gen}\left(\begin{bmatrix} 1 \\ i \end{bmatrix}\right); \lambda = 1 - i, E_{1-i} = \text{gen}\left(\begin{bmatrix} 1 \\ -i \end{bmatrix}\right)$

29. $\lambda = 1 + i, E_{1+i} = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right); \lambda = 1 - i, E_{1-i} = \text{gen}\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$

31. $\lambda = 1, 2$ **33.** $\lambda = 4$

Ejercicios 4.2

1. 16 **3.** 0 **5.** -18 **7.** 6
9. -12 **11.** $a^2b + ab^2$ **13.** 4 **15.** $abdg$
17. 0 **25.** 2 **27.** -24 **29.** 0
31. 0 **33.** -24 **35.** 8 **37.** -4
39. -8 **45.** $k \neq 0, 2$ **47.** -6 **49.** $-\frac{3}{2}$
51. $(-2)3^n$
53. $\det(AB) = (\det A)(\det B) = (\det B)(\det A) = \det(BA)$
55. 0, 1 **57.** $x = \frac{3}{2}, y = -\frac{1}{2}$
59. $x = -1, y = 0, z = 1$ **61.** $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$
63. $\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$

Ejercicios 4.3

1. **(a)** $\lambda^2 - 7\lambda + 12$ **(b)** $\lambda = 3, 4$
(c) $E_3 = \text{gen}\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}\right); E_4 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$
(d) Las multiplicidades algebraica y geométrica son todas 1.

3. **(a)** $-\lambda^3 + 2\lambda^2 + 5\lambda - 6$
(b) $\lambda = -2, 1, 3$
(c) $E_{-2} = \text{gen}\left(\begin{bmatrix} 1 \\ -3 \\ 0 \end{bmatrix}\right); E_1 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right);$
$E_3 = \text{gen}\left(\begin{bmatrix} 1 \\ 2 \\ 10 \end{bmatrix}\right)$
(d) Las multiplicidades algebraica y geométrica son todas 1.

5. **(a)** $-\lambda^3 + \lambda^2$ **(b)** $\lambda = 0, 1$
(c) $E_0 = \text{gen}\left(\begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}\right); E_1 = \text{gen}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right)$
(d) $\lambda = 0$ tiene multiplicidad algebraica 2 y multiplicidad geométrica 1; $\lambda = 1$ tiene multiplicidades algebraica y geométrica 1.

7. **(a)** $-\lambda^3 + 9\lambda^2 - 27\lambda + 27$
(b) $\lambda = 3$
(c) $E_3 = \text{gen}\left(\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right)$
(d) $\lambda = 3$ tiene multiplicidad algebraica 3 y multiplicidad geométrica 2.

9. **(a)** $\lambda^4 - 6\lambda^3 + 9\lambda^2 + 4\lambda - 12$
(b) $\lambda = -1, 2, 3$
(c) $E_{-1} = \text{gen}\left(\begin{bmatrix} 0 \\ 0 \\ -2 \\ 1 \end{bmatrix}\right); E_2 = \text{gen}\left(\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}\right);$
$E_3 = \text{gen}\left(\begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}\right)$
(d) $\lambda = -1$ y $\lambda = 3$ tienen multiplicidades algebraica y geométrica 1; $\lambda = 2$ tiene multiplicidad algebraica 2 y multiplicidad geométrica 1.

11. **(a)** $\lambda^4 - 4\lambda^3 + 2\lambda^2 + 4\lambda - 3$

(b) $\lambda = -1, 1, 3$

(c) $E_{-1} = \text{gen}\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right)$;

$E_1 = \text{gen}\left(\begin{bmatrix} -2 \\ 0 \\ 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 0 \\ 3 \end{bmatrix}\right)$;

$E_3 = \text{gen}\left(\begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix}\right)$

(d) $\lambda = -1$ y $\lambda = 3$ tienen multiplicidades algebraica y geométrica 1; $\lambda = 1$ tiene multiplicidades algebraica y geométrica 2.

15. $\begin{bmatrix} 2^{-9} + 3 \cdot 2^{10} \\ -2^{-9} + 3 \cdot 2^{10} \end{bmatrix}$ **17.** $\begin{bmatrix} 2 \\ (2 \cdot 3^{20} - 1)/3^{20} \\ 2 \end{bmatrix}$

23. **(a)** $\lambda = -2, E_{-2} = \text{gen}\left(\begin{bmatrix} 2 \\ -5 \end{bmatrix}\right); \lambda = 5, E_5 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$

(b) **(i)** $\lambda = -\frac{1}{2}, E_{-1/2} = \text{gen}\left(\begin{bmatrix} 2 \\ -5 \end{bmatrix}\right); \lambda = \frac{1}{5}, E_{1/5} = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$

(iii) $\lambda = 0, E_0 = \text{gen}\left(\begin{bmatrix} 2 \\ -5 \end{bmatrix}\right); \lambda = 7, E_7 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$

27. $\begin{bmatrix} -3 & 4 & -12 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, -\lambda^3 - 3\lambda^2 + 4\lambda - 12$

35. $A^2 = 4A - 5I, A^3 = 11A - 20I$

$A^4 = 24A - 55I$

37. $A^{-1} = -\frac{1}{5}A + \frac{4}{5}I, A^{-2} = -\frac{4}{25}A + \frac{11}{25}I$

Ejercicios 4.4

1. El polinomio característico de A es $\lambda^2 - 5\lambda + 1$, pero el de B es $\lambda^2 - 2\lambda + 1$.

3. Los eigenvalores de A son $\lambda = 2$ y $\lambda = 4$, pero los de B son $\lambda = 1$ y $\lambda = 4$.

5. $\lambda_1 = 4, E_4 = \text{gen}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right); \lambda_2 = 3, E_3 = \text{gen}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$

7. $\lambda_1 = 6, E_6 = \text{gen}\left(\begin{bmatrix} 3 \\ 2 \\ 3 \end{bmatrix}\right); \lambda_2 = -2, E_{-2} = \text{gen}\left(\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right)$

9. No es diagonalizable

11. $P = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & -2 & 0 \end{bmatrix}, D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

13. No es diagonalizable

15. $P = \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$

17. $\begin{bmatrix} 35839 & -69630 \\ -11605 & 24234 \end{bmatrix}$

19. $\begin{bmatrix} (3^k + 3(-1)^k)/4 & (3^{k+1} - 3(-1)^k)/4 \\ (3^k - (-1)^k)/4 & (3^{k+1} + (-1)^k)/4 \end{bmatrix}$

21. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

23.
$$\begin{bmatrix} (5 + 2^{k+2} + (-3)^k)/10 & (2^k - (-3)^k)/5 & (-5 + 2^{k+2} + (-3)^k)/10 \\ (2^{k+1} - 2(-3)^k)/5 & (2^k + 4(-3)^k)/5 & (2^{k+1} - 2(-3)^k)/5 \\ (-5 + 2^{k+2} + (-3)^k)/10 & (2^k - (-3)^k)/5 & (5 + 2^{k+2} + (-3)^k)/10 \end{bmatrix}$$

25. $k = 0$ **27.** $k = 0$

29. Todos los valores reales de k

35. If $A \sim B$, entonces existe una matriz invertible P tal que $B = P^{-1}AP$. Por tanto, se tiene

$$\text{tr}(B) = \text{tr}(P^{-1}AP) = \text{tr}(P^{-1}(AP)) = \text{tr}((AP)P^{-1})$$
$$= \text{tr}(APP^{-1}) = \text{tr}(AI) = \text{tr}(A)$$

al usar el ejercicio 45 de la sección 3.2.

37. $P = \begin{bmatrix} 7 & -2 \\ 10 & -3 \end{bmatrix}$

39. $P = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{3}{2} & -\frac{3}{2} & 1 \\ -\frac{5}{2} & -\frac{3}{2} & 0 \end{bmatrix}$

49. (b) $\dim E_{-1} = 1, \dim E_1 = 2, \dim E_2 = 3$

Ejercicios 4.5

1. (a) $\begin{bmatrix} 1 \\ 2.5 \end{bmatrix}$, 6.000

(b) $\lambda_1 = 6$

3. (a) $\begin{bmatrix} 1 \\ 0.618 \end{bmatrix}$, 2.618

(b) $\lambda_1 = (3 + \sqrt{5})/2 \approx 2.618$

5. (a) $m_5 = 11.001, \mathbf{y}_5 = \begin{bmatrix} -0.333 \\ 1.000 \end{bmatrix}$

7. (a) $m_8 = 10.000, \mathbf{y}_8 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

9.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 26 \\ 8 \end{bmatrix}$	$\begin{bmatrix} 17.692 \\ 5.923 \end{bmatrix}$	$\begin{bmatrix} 18.018 \\ 6.004 \end{bmatrix}$	$\begin{bmatrix} 17.999 \\ 6.000 \end{bmatrix}$	$\begin{bmatrix} 18.000 \\ 6.000 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.308 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.335 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.333 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.333 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.333 \end{bmatrix}$
m_k	1	26	17.692	18.018	17.999	18.000

Por tanto, $\lambda_1 \approx 18, \mathbf{v}_1 \approx \begin{bmatrix} 1 \\ 0.333 \end{bmatrix}$.

11.

k	0	1	2	3	4	5	6
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 7 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 7.571 \\ 2.857 \end{bmatrix}$	$\begin{bmatrix} 7.755 \\ 3.132 \end{bmatrix}$	$\begin{bmatrix} 7.808 \\ 3.212 \end{bmatrix}$	$\begin{bmatrix} 7.823 \\ 3.234 \end{bmatrix}$	$\begin{bmatrix} 7.827 \\ 3.240 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.286 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.377 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.404 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.411 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.413 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.414 \end{bmatrix}$
m_k	1	7	7.571	7.755	7.808	7.823	7.827

Por tanto, $\lambda_1 \approx 7.827, \mathbf{v}_1 \approx \begin{bmatrix} 1 \\ 0.414 \end{bmatrix}$.

13.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 21 \\ 15 \\ 13 \end{bmatrix}$	$\begin{bmatrix} 16.809 \\ 12.238 \\ 10.714 \end{bmatrix}$	$\begin{bmatrix} 17.011 \\ 12.371 \\ 10.824 \end{bmatrix}$	$\begin{bmatrix} 16.999 \\ 12.363 \\ 10.818 \end{bmatrix}$	$\begin{bmatrix} 17.000 \\ 12.363 \\ 10.818 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.714 \\ 0.619 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.728 \\ 0.637 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.727 \\ 0.636 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.727 \\ 0.636 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.727 \\ 0.636 \end{bmatrix}$
m_k	1	21	16.809	17.011	16.999	17.000

Por tanto, $\lambda_1 \approx 17, \mathbf{v}_1 \approx \begin{bmatrix} 1 \\ 0.727 \\ 0.636 \end{bmatrix}$.

15. $\lambda_1 \approx 5, \mathbf{v}_1 \approx \begin{bmatrix} 1 \\ 0 \\ 0.333 \end{bmatrix}$

17.

k	0	1	2	3	4	5	6
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 7 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 7.571 \\ 2.857 \end{bmatrix}$	$\begin{bmatrix} 7.755 \\ 3.132 \end{bmatrix}$	$\begin{bmatrix} 7.808 \\ 3.212 \end{bmatrix}$	$\begin{bmatrix} 7.823 \\ 3.234 \end{bmatrix}$	$\begin{bmatrix} 7.827 \\ 3.240 \end{bmatrix}$
$R(\mathbf{x}_k)$	7	7.755	7.823	7.828	7.828	7.828	7.828
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.286 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.377 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.404 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.411 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.413 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.414 \end{bmatrix}$

19.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 21 \\ 15 \\ 13 \end{bmatrix}$	$\begin{bmatrix} 16.809 \\ 12.238 \\ 10.714 \end{bmatrix}$	$\begin{bmatrix} 17.011 \\ 12.371 \\ 10.824 \end{bmatrix}$	$\begin{bmatrix} 16.999 \\ 12.363 \\ 10.818 \end{bmatrix}$	$\begin{bmatrix} 17.000 \\ 12.363 \\ 10.818 \end{bmatrix}$
$R(\mathbf{x}_k)$	16.333	16.998	17.000	17.000	17.000	17.000
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.714 \\ 0.619 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.728 \\ 0.637 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.727 \\ 0.636 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.727 \\ 0.636 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.727 \\ 0.636 \end{bmatrix}$

21.

k	0	1	2	3	4	5	6	7	8
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 5 \\ 4 \end{bmatrix}$	$\begin{bmatrix} 4.8 \\ 3.2 \end{bmatrix}$	$\begin{bmatrix} 4.667 \\ 2.667 \end{bmatrix}$	$\begin{bmatrix} 4.571 \\ 2.286 \end{bmatrix}$	$\begin{bmatrix} 4.500 \\ 2.000 \end{bmatrix}$	$\begin{bmatrix} 4.444 \\ 1.778 \end{bmatrix}$	$\begin{bmatrix} 4.400 \\ 1.600 \end{bmatrix}$	$\begin{bmatrix} 4.364 \\ 1.455 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.8 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.667 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.571 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.500 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.444 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.400 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.364 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.333 \end{bmatrix}$
m_k	1	5	4.8	4.667	4.571	4.500	4.444	4.400	4.364

Puesto que $\lambda_1 = \lambda_2 = 4, \mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, m_k converge lentamente en la respuesta exacta.

23.

k	0	1	2	3	4	5	6	7	8
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 4.2 \\ 3.2 \\ 0.2 \end{bmatrix}$	$\begin{bmatrix} 4.048 \\ 3.048 \\ 0.048 \end{bmatrix}$	$\begin{bmatrix} 4.012 \\ 3.012 \\ 0.012 \end{bmatrix}$	$\begin{bmatrix} 4.003 \\ 3.003 \\ 0.003 \end{bmatrix}$	$\begin{bmatrix} 4.001 \\ 3.001 \\ 0.001 \end{bmatrix}$	$\begin{bmatrix} 4.000 \\ 3.000 \\ 0.000 \end{bmatrix}$	$\begin{bmatrix} 4.000 \\ 3.000 \\ 0.000 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.8 \\ 0.2 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.762 \\ 0.048 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.753 \\ 0.012 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.751 \\ 0.003 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.750 \\ 0.001 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.750 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.750 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0.750 \\ 0 \end{bmatrix}$
m_k	1	5	4.2	4.048	4.012	4.003	4.001	4.000	4.000

En este caso, $\lambda_1 = \lambda_2 = 4$ y $E_4 = \text{gen}\left(\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right)$.

Claramente, m_k converge en 4 y $\mathbf{y}_k$ converge en un vector en el eigenespacio E_4, a saber, $\begin{bmatrix}1\\0\\0\end{bmatrix} + 0.75\begin{bmatrix}0\\1\\0\end{bmatrix}$.

25.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}-1\\-1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}-1\\-1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$	$\begin{bmatrix}1\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\end{bmatrix}$
m_k	1	1	-1	1	-1	1

Los eigenvalores exactos son complejos (i y $-i$), de modo que el método de potencias posiblemente no pueda convergir en el eigenvalor dominante o en el eigenvector dominante si comienza con una iteración inicial *real*. En vez de ello, el método de potencia oscila entre dos conjuntos de vectores reales.

27.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}3\\4\\3\end{bmatrix}$	$\begin{bmatrix}2.500\\4.000\\2.500\end{bmatrix}$	$\begin{bmatrix}2.250\\4.000\\2.250\end{bmatrix}$	$\begin{bmatrix}2.125\\4.000\\2.125\end{bmatrix}$	$\begin{bmatrix}2.063\\4.000\\2.063\end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}0.750\\1\\0.750\end{bmatrix}$	$\begin{bmatrix}0.625\\1\\0.625\end{bmatrix}$	$\begin{bmatrix}0.562\\1\\0.562\end{bmatrix}$	$\begin{bmatrix}0.531\\1\\0.531\end{bmatrix}$	$\begin{bmatrix}0.516\\1\\0.516\end{bmatrix}$
m_k	1	4	4	4	4	4

Los eigenvalores son $\lambda_1 = -12$, $\lambda_2 = 4$, $\lambda_3 = 2$, con eigenvectores correspondientes

$$\mathbf{v}_1 = \begin{bmatrix}1\\0\\-1\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix}1\\2\\1\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix}1\\0\\1\end{bmatrix}.$$

Puesto que $\mathbf{x}_0 = \frac{1}{2}\mathbf{v}_2 + \frac{1}{2}\mathbf{v}_3$, el vector inicial $\mathbf{x}_0$ tiene un componente cero en la dirección del eigenvector dominante, de modo que el método de potencia no puede convergir en el eigenvalor/eigenvector dominante. En vez de ello, converge en un *segundo* par eigenvalor/eigenvector, como muestran los cálculos.

29. Aplique el método de potencia a $A - 18I = \begin{bmatrix} -4 & 12 \\ 5 & -15 \end{bmatrix}$.

k	0	1	2	3
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 8 \\ -10 \end{bmatrix}$	$\begin{bmatrix} 15.2 \\ -19 \end{bmatrix}$	$\begin{bmatrix} 15.2 \\ -19 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.8 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.8 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.8 \\ 1 \end{bmatrix}$
m_k	1	-10	-19	-19

Por ende, -19 es el eigenvalor dominante de $A - 18I$, y $\lambda_2 = -19 + 18 = -1$ es el segundo eigenvalor de A.

31. Aplique el método de potencia a $A - 17I = \begin{bmatrix} -8 & 4 & 8 \\ 4 & -2 & -4 \\ 8 & -4 & -8 \end{bmatrix}$.

k	0	1	2	3
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 4 \\ -2 \\ -4 \end{bmatrix}$	$\begin{bmatrix} -18 \\ 9 \\ 18 \end{bmatrix}$	$\begin{bmatrix} -18 \\ 9 \\ 18 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -0.5 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -0.5 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -0.5 \\ -1 \end{bmatrix}$
m_k	1	4	-18	-18
$R(\mathbf{x}_k)$	-0.667	-18	-18	-18

En este caso, no existe eigenvalor dominante. (Podría elegir 18 o -18 para m_k, $k \geq 2$.) Sin embargo, el método de cociente de Rayleigh (ejercicios 17-20) converge en -18. Por ende, -18 es el eigenvalor dominante de $A - 17\text{I}$ y $\lambda_2 = -18 + 17 = -1$ es el segundo eigenvalor de A.

33.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ -0.5 \end{bmatrix}$	$\begin{bmatrix} -0.833 \\ 1.056 \end{bmatrix}$	$\begin{bmatrix} 0.798 \\ -0.997 \end{bmatrix}$	$\begin{bmatrix} 0.800 \\ -1.000 \end{bmatrix}$	$\begin{bmatrix} 0.800 \\ -1.000 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} -0.789 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.801 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.800 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.800 \\ 1 \end{bmatrix}$
m_k	1	0.5	1.056	-0.997	-1.000	-1.000

Por tanto, el eigenvalor de A que es más pequeño en magnitud es $1/(-1) = -1$.

35.

k	0	1	2	3	4	5
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} -0.500 \\ 0.000 \\ 0.500 \end{bmatrix}$	$\begin{bmatrix} 0.500 \\ 0.333 \\ -0.500 \end{bmatrix}$	$\begin{bmatrix} 0.500 \\ 0.111 \\ -0.500 \end{bmatrix}$	$\begin{bmatrix} 0.500 \\ 0.259 \\ -0.500 \end{bmatrix}$	$\begin{bmatrix} 0.500 \\ 0.160 \\ -0.500 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} -1.000 \\ 0.000 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} -1.000 \\ -0.667 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} -1.000 \\ -0.222 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} -1.000 \\ -0.518 \\ 1.000 \end{bmatrix}$	$\begin{bmatrix} -1.000 \\ -0.321 \\ 1.000 \end{bmatrix}$
m_k	1	-0.500	-0.500	-0.500	-0.500	-0.500

Claramente, m_k converge en -0.5, de modo que el eigenvalor más pequeño de A es $1/(-0.5) = -2$.

37. Los cálculos son los mismos que los del ejercicio 33.

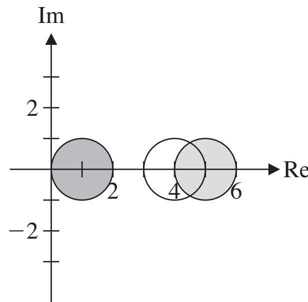
39. Aplique el método de potencia inverso a $A - 5I = \begin{bmatrix} -1 & 0 & 6 \\ -1 & -2 & 1 \\ 6 & 0 & -1 \end{bmatrix}$. Al tomar $\mathbf{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, se tiene

k	0	1	2	3
$\mathbf{x}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0.200 \\ -0.500 \\ 0.200 \end{bmatrix}$	$\begin{bmatrix} -0.080 \\ -0.500 \\ -0.080 \end{bmatrix}$	$\begin{bmatrix} 0.032 \\ -0.500 \\ 0.032 \end{bmatrix}$
$\mathbf{y}_k$	$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -0.400 \\ 1 \\ -0.400 \end{bmatrix}$	$\begin{bmatrix} 0.160 \\ 1 \\ 0.160 \end{bmatrix}$	$\begin{bmatrix} -0.064 \\ 1 \\ -0.064 \end{bmatrix}$
m_k	1	-0.500	-0.500	-0.500

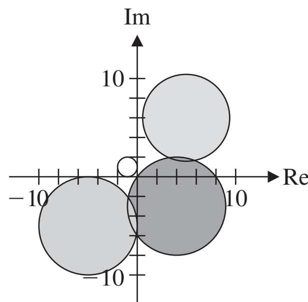
Claramente, m_k converge en -0.5, de modo que el eigenvalor de A más cercano a 5 es $5 + 1/(-0.5) = 5 - 2 = 3$.

41. 0.732 **43.** -0.619

47.



49.



51. *Sugerencia:* demuestre que 0 no está contenido en algún disco de Gerschgorin y luego aplique el Teorema 4.16.

53. El ejercicio 52 implica que $|\lambda|$ es menor o igual a todas las sumas de columna de A para todo eigenvalor λ. Pero, para una matriz estocástica, todas las sumas de columna son 1. Por tanto $|\lambda| \leq 1$.

Ejercicios 4.6

1. No es regular **3.** Regular

5. No es regular **7.** $L = \begin{bmatrix} \frac{1}{5} & \frac{1}{5} \\ \frac{4}{5} & \frac{4}{5} \end{bmatrix}$

9. $L = \begin{bmatrix} 0.304 & 0.304 & 0.304 \\ 0.354 & 0.354 & 0.354 \\ 0.342 & 0.342 & 0.342 \end{bmatrix}$

11. $1, \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ **13.** $2, \begin{bmatrix} 16 \\ 4 \\ 1 \end{bmatrix}$

15. La población aumenta, disminuye y es constante, respectivamente.

17.
$$P^{-1}LP = \begin{bmatrix} b_1 & b_2s_1 & b_3s_1s_2 & \cdots & b_{n-1}s_1s_2\cdots s_{n-2} & b_ns_1s_2\cdots s_{n-1} \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.$$

El polinomio característico de L es $(\lambda^n - b_1\lambda^{n-1} - b_2s_1\lambda^{n-2} - b_3s_1s_2\lambda^{n-3} - \cdots - b_ns_1s_2\cdots s_{n-1})(-1)^n$.

19. $\lambda \approx 1.746$, $\mathbf{p} \approx \begin{bmatrix} 0.660 \\ 0.264 \\ 0.076 \end{bmatrix}$

21. $\lambda \approx 1.092$, $\mathbf{p} \approx \begin{bmatrix} 0.535 \\ 0.147 \\ 0.094 \\ 0.078 \\ 0.064 \\ 0.053 \\ 0.029 \end{bmatrix}$

25. (a) $h \approx 0.082$ **29.** $3, \begin{bmatrix} \frac{3}{5} \\ \frac{2}{5} \end{bmatrix}$

31. $3, \begin{bmatrix} \frac{1}{2} \\ \frac{1}{4} \\ \frac{1}{4} \end{bmatrix}$ **33.** Reducible

35. Irreductible **43.** 1, 2, 4, 8, 16

45. 0, 1, 1, 0, -1 **47.** $x_n = 4^n - (-1)^n$

49. $y_n = (n - \frac{1}{2})2^n$

51. $b_n = \dfrac{1}{2\sqrt{3}}[(1 + \sqrt{3})^n - (1 - \sqrt{3})^n]$

57. **(a)** $d_1 = 1, d_2 = 2, d_3 = 3, d_4 = 5, d_5 = 8$
(b) $d_n = d_{n-1} + d_{n-2}$
(c) $d_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right]$

59. La solución general es $x(t) = -3C_1e^{-t} + C_2e^{4t}, y(t) = 2C_1e^{-t} + C_2e^{4t}$. La solución específica es $x(t) = -3e^{-t} + 3e^{4t}, y(t) = 2e^{-t} + 3e^{4t}$.

61. La solución general es $x_1(t) = (1 + \sqrt{2})C_1e^{\sqrt{2}t} + (1 - \sqrt{2})C_2e^{-\sqrt{2}t}, x_2(t) = C_1e^{\sqrt{2}t} + C_2e^{-\sqrt{2}t}$. La solución específica es $x_1(t) = (2 + \sqrt{2})e^{\sqrt{2}t}/4 + (2 - \sqrt{2})e^{-\sqrt{2}t}/4, x_2(t) = \sqrt{2}e^{\sqrt{2}t}/4 - \sqrt{2}e^{-\sqrt{2}t}/4$.

63. La solución general es $x(t) = -C_1 + C_3e^{-t}, y(t) = C_1 + C_2e^{t} - C_3e^{-t}, z(t) = C_1 + C_2e^{t}$. La solución específica es $x(t) = 2 - e^{-t}, y(t) = -2 + e^{t} + e^{-t}$, $z(t) = -2 + e^{t}$.

65. **(a)** $x(t) = -120e^{8t/5} + 520e^{11t/10}, y(t) = 240e^{8t/5} + 260^{11t/10}$. La cepa X muere después de aproximadamente 2.93 días; la cepa Y sigue creciendo.

67. $a = 10, b = 20; x(t) = 10e^{t}(\cos t + \operatorname{sen} t) + 10$, $y(t) = 10e^{t}(\cos t - \operatorname{sen} t) + 20$. La especie Y muere cuando $t \approx 1.22$.

71. $x(t) = C_1e^{2t} + C_2e^{3t}$

77. **(a)** $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}3\\3\end{bmatrix}, \begin{bmatrix}9\\9\end{bmatrix}, \begin{bmatrix}27\\27\end{bmatrix}$ **(c)** Repelente

79. **(a)** $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\1\end{bmatrix}$ **(c)** Ninguno

81. **(a)** $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}0.5\\-1\end{bmatrix}, \begin{bmatrix}1.75\\-0.5\end{bmatrix}, \begin{bmatrix}3.125\\-1.75\end{bmatrix}$ **(c)** Punto de silla

83. **(a)** $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}0.6\\0.6\end{bmatrix}, \begin{bmatrix}0.36\\0.36\end{bmatrix}, \begin{bmatrix}0.216\\0.216\end{bmatrix}$ **(c)** Atractor

85. $r = \sqrt{2}, \theta = 45°$, repelente espiral

87. $r = 2, \theta = -60°$, repelente espiral

89. $P = \begin{bmatrix}-1 & -1\\1 & 0\end{bmatrix}, C = \begin{bmatrix}0.2 & -0.1\\0.1 & 0.2\end{bmatrix}$, atractor espiral

91. $P = \begin{bmatrix}1/2 & -\sqrt{3}/2\\1 & 0\end{bmatrix}, C = \begin{bmatrix}1/2 & -\sqrt{3}/2\\\sqrt{3}/2 & 1/2\end{bmatrix}$, centro orbital

Preguntas de repaso

1. **(a)** F **(c)** F **(e)** F **(g)** V **(i)** F

3. -18

5. Puesto que $A^T = -A$, se tiene $\det A = \det(A^T) = \det(-A) = (-1)^n \det A = -\det A$ por el Teorema 4.7 y el hecho de que n es impar. Se tiene que $\det A = 0$.

7. $A\mathbf{x} = \begin{bmatrix}5\\10\end{bmatrix} = 5\mathbf{x}, \lambda = 5$

9. **(a)** $4 - 3\lambda^2 - \lambda^3$

(c) $E_1 = \text{gen}\left(\begin{bmatrix}1\\-1\\0\end{bmatrix}\right), E_{-2} = \text{gen}\left(\begin{bmatrix}2\\-1\\0\end{bmatrix}, \begin{bmatrix}1\\0\\1\end{bmatrix}\right)$

11. $\begin{bmatrix}162\\158\end{bmatrix}$ **13.** No es similar **15.** No es similar

17. 0, 1, o -1

19. Si $A\mathbf{x} = \lambda\mathbf{x}$, entonces $(A^2 - 5A + 2I)\mathbf{x} = A^2\mathbf{x} - 5A\mathbf{x} + 2\mathbf{x} = 3^2\mathbf{x} - 5(3\mathbf{x}) + 2\mathbf{x} = -4\mathbf{x}$.

Capítulo 5

Ejercicios 5.1

1. Ortogonal **3.** No es ortogonal **5.** Ortogonal

7. $[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix}\frac{1}{2}\\-1\end{bmatrix}$ **9.** $[\mathbf{w}]_{\mathcal{B}} = \begin{bmatrix}0\\\frac{2}{3}\\\frac{1}{3}\end{bmatrix}$ **11.** Ortonormal

13. $\begin{bmatrix}1/3\\2/3\\2/3\end{bmatrix}, \begin{bmatrix}2/\sqrt{5}\\-1/\sqrt{5}\\0\end{bmatrix}, \begin{bmatrix}2/3\sqrt{5}\\4/3\sqrt{5}\\-5/3\sqrt{5}\end{bmatrix}$

15. Ortonormal

17. Ortogonal, $\begin{bmatrix}1/\sqrt{2} & -1/\sqrt{2}\\1/\sqrt{2} & 1/\sqrt{2}\end{bmatrix}$

19. Ortogonal, $\begin{bmatrix}\cos\theta \operatorname{sen}\theta & \cos^2\theta & \operatorname{sen}\theta\\-\cos\theta & \operatorname{sen}\theta & 0\\-\operatorname{sen}^2\theta & -\cos\theta\operatorname{sen}\theta & \cos\theta\end{bmatrix}$

21. No es ortogonal

27. $$\cos(\angle(Q\mathbf{x}, Q\mathbf{y})) = \frac{(Q\mathbf{x})\cdot(Q\mathbf{y})}{\|Q\mathbf{x}\|\|Q\mathbf{y}\|} = \frac{(Q\mathbf{x})^TQ\mathbf{y}}{\sqrt{(Q\mathbf{x})^TQ\mathbf{x}}\sqrt{(Q\mathbf{y})^TQ\mathbf{y}}} = \frac{\mathbf{x}^TQ^TQ\mathbf{y}}{\sqrt{\mathbf{x}^TQ^TQ\mathbf{x}}\sqrt{\mathbf{y}^TQ^TQ\mathbf{y}}} = \frac{\mathbf{x}^T\mathbf{y}}{\sqrt{\mathbf{x}^T\mathbf{x}}\sqrt{\mathbf{y}^T\mathbf{y}}} = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|} = \cos(\angle(\mathbf{x}, \mathbf{y}))$$

29. Rotación, $\theta = 45°$ **31.** Reflexión, $y = \sqrt{3}x$

33. (a) $A(A^T + B^T)B = AA^TB + AB^TB = IB + AI = B + A = A + B$

(b) Del inciso (a),

$$\begin{aligned}\det(A + B) &= \det(A(A^T + B^T)B)\\ &= \det A \det(A^T + B^T)\det B\\ &= \det A \det((A + B)^T)\det B\\ &= \det A \det(A + B)\det B\end{aligned}$$

Suponga que $\det A + \det B = 0$ (de modo que $\det B = -\det A$) pero que $A + B$ es invertible. Entonces $\det(A + B) \neq 0$, de modo que $1 = \det A \det B = \det A(-\det A) = -(\det A)^2$. Esto es imposible, así que se concluye que $A + B$ no puede ser invertible.

Ejercicios 5.2

1. $W^\perp = \left\{\begin{bmatrix} x\\ y\end{bmatrix} : x + 2y = 0\right\}, \mathcal{B}^\perp = \left\{\begin{bmatrix} -2\\ 1\end{bmatrix}\right\}$

3. $W^\perp = \left\{\begin{bmatrix} x\\ y\\ z\end{bmatrix} : x = t, y = t, z = -t\right\}, \mathcal{B}^\perp = \left\{\begin{bmatrix} 1\\ 1\\ -1\end{bmatrix}\right\}$

5. $W^\perp = \left\{\begin{bmatrix} x\\ y\\ z\end{bmatrix} : x - y + 3z = 0\right\}, \mathcal{B}^\perp = \left\{\begin{bmatrix} 1\\ 1\\ 0\end{bmatrix}, \begin{bmatrix} 0\\ 3\\ 1\end{bmatrix}\right\}$

7. renglón(A):$\{[1\ \ 0\ \ 1], [0\ \ 1\ \ -2]\}$, nulo$(A)$: $\left\{\begin{bmatrix} 1\\ 2\\ 1\end{bmatrix}\right\}$

9. col(A): $\left\{\begin{bmatrix} 1\\ 5\\ 0\\ -1\end{bmatrix}, \begin{bmatrix} -1\\ 2\\ 1\\ -1\end{bmatrix}\right\}$, nulo$(A^T)$: $\left\{\begin{bmatrix} 1\\ 0\\ 2\\ 1\end{bmatrix}, \begin{bmatrix} -5\\ 1\\ -7\\ 0\end{bmatrix}\right\}$

11. $\left\{\begin{bmatrix} 1\\ -10\\ -4\end{bmatrix}\right\}$ **13.** $\left\{\begin{bmatrix} -4\\ 1\\ 0\\ 3\end{bmatrix}, \begin{bmatrix} -3\\ 0\\ 1\\ 0\end{bmatrix}\right\}$

15. $\begin{bmatrix} \frac{3}{2}\\ \frac{3}{2}\end{bmatrix}$ **17.** $\begin{bmatrix} -\frac{1}{2}\\ \frac{1}{2}\\ 3\end{bmatrix}$

19. $\mathbf{v} = \begin{bmatrix} -\frac{2}{5}\\ -\frac{6}{5}\end{bmatrix} + \begin{bmatrix} \frac{12}{5}\\ -\frac{4}{5}\end{bmatrix}$ **21.** $\mathbf{v} = \begin{bmatrix} \frac{7}{2}\\ -2\\ \frac{7}{2}\end{bmatrix} + \begin{bmatrix} \frac{1}{2}\\ 0\\ -\frac{1}{2}\end{bmatrix}$

25. No

Ejercicios 5.3

1. $\mathbf{v}_1 = \begin{bmatrix} 1\\ 1\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} -1/2\\ 1/2\end{bmatrix}; \mathbf{q}_1 = \begin{bmatrix} 1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} -1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$

3. $\mathbf{v}_1 = \begin{bmatrix} 1\\ -1\\ -1\end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2\\ 1\\ 1\end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0\\ -1\\ 1\end{bmatrix}; \mathbf{q}_1 = \begin{bmatrix} 1/\sqrt{3}\\ -1/\sqrt{3}\\ -1/\sqrt{3}\end{bmatrix},$

$\mathbf{q}_2 = \begin{bmatrix} 2/\sqrt{6}\\ 1/\sqrt{6}\\ 1/\sqrt{6}\end{bmatrix}, \mathbf{q}_3 = \begin{bmatrix} 0\\ -1/\sqrt{2}\\ 1/\sqrt{2}\end{bmatrix}$

5. $\left\{\begin{bmatrix} 1\\ 1\\ 0\end{bmatrix}, \begin{bmatrix} -\frac{1}{2}\\ \frac{1}{2}\\ 2\end{bmatrix}\right\}$ **7.** $\mathbf{v} = \begin{bmatrix} -\frac{2}{9}\\ \frac{2}{9}\\ \frac{8}{9}\end{bmatrix} + \begin{bmatrix} \frac{38}{9}\\ -\frac{38}{9}\\ \frac{19}{9}\end{bmatrix}$

9. $\left\{\begin{bmatrix} 0\\ 1\\ 1\end{bmatrix}, \begin{bmatrix} 1\\ -\frac{1}{2}\\ \frac{1}{2}\end{bmatrix}, \begin{bmatrix} \frac{2}{3}\\ \frac{2}{3}\\ -\frac{2}{3}\end{bmatrix}\right\}$

11. $\left\{\begin{bmatrix} 3\\ 1\\ 5\end{bmatrix}, \begin{bmatrix} -\frac{3}{35}\\ \frac{34}{35}\\ -\frac{1}{7}\end{bmatrix}, \begin{bmatrix} -\frac{15}{34}\\ 0\\ \frac{9}{34}\end{bmatrix}\right\}$

13. $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6}\\ 0 & 1/\sqrt{3} & -2/\sqrt{6}\\ -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6}\end{bmatrix}$

15. $\begin{bmatrix} 0 & 2/\sqrt{6} & 1/\sqrt{3}\\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3}\end{bmatrix}\begin{bmatrix} \sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2}\\ 0 & 3/\sqrt{6} & 1/\sqrt{6}\\ 0 & 0 & 2/\sqrt{3}\end{bmatrix}$

17. $R = \begin{bmatrix} 3 & 9 & \frac{1}{3}\\ 0 & 6 & \frac{2}{3}\\ 0 & 0 & \frac{7}{3}\end{bmatrix}$

19. $A = AI$

21. $A^{-1} = (QR)^{-1} = R^{-1}Q^{-1} = R^{-1}Q^T =$

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{6} & -1/2\sqrt{3}\\ 0 & 2/\sqrt{6} & -1/2\sqrt{3}\\ 0 & 0 & 3/2\sqrt{3}\end{bmatrix}\begin{bmatrix} 0 & 1/\sqrt{2} & 1/\sqrt{2}\\ 2/\sqrt{6} & -1/\sqrt{6} & 1/\sqrt{6}\\ 1/\sqrt{3} & 1/\sqrt{3} & -1/\sqrt{3}\end{bmatrix}$$

23. Sea $R\mathbf{x} = \mathbf{0}$. Entonces $A\mathbf{x} = QR\mathbf{x} = Q\mathbf{0} = \mathbf{0}$. Puesto que $A\mathbf{x}$ representa una combinación lineal de las columnas de A (que son linealmente independientes), debe tener $\mathbf{x} = \mathbf{0}$. Por tanto, R es invertible, por el teorema fundamental.

Ejercicios 5.4

1. $Q = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2}\\ 1/\sqrt{2} & -1/\sqrt{2}\end{bmatrix}, D = \begin{bmatrix} 5 & 0\\ 0 & 3\end{bmatrix}$

3. $Q = \begin{bmatrix} 2/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{3} & -2/\sqrt{6} \end{bmatrix}, D = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$

5. $Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}, D = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -2 \end{bmatrix}$

7. $Q = \begin{bmatrix} -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix}, D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

9. $Q = \begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix},$

$D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

11. $Q^TAQ = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} a & b \\ b & a \end{bmatrix} \cdot$

$\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} a+b & 0 \\ 0 & a-b \end{bmatrix} = D$

13. (a) Si A y B son diagonalizables ortogonalmente, entonces cada una es simétrica, por el teorema espectral. Por tanto, $A + B$ es simétrica, por el ejercicio 35 de la sección 3.2, y por tanto es diagonalizable ortogonalmente, por el teorema espectral.

15. Si A y B son diagonalizables ortogonalmente, entonces cada una es simétrica, por el teorema espectral. Puesto que $AB = BA$, AB también es simétrica, por el ejercicio 36 de la sección 3.2. Por tanto, AB es diagonalizable ortogonalmente, por el teorema espectral.

17. $A = \begin{bmatrix} \frac{5}{2} & \frac{5}{2} \\ \frac{5}{2} & \frac{5}{2} \end{bmatrix} + \begin{bmatrix} \frac{3}{2} & -\frac{3}{2} \\ -\frac{3}{2} & \frac{3}{2} \end{bmatrix}$

19. $A = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 2 & 2 \\ 0 & 2 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$

21. $\begin{bmatrix} \frac{1}{2} & -\frac{3}{2} \\ -\frac{3}{2} & \frac{1}{2} \end{bmatrix}$ **23.** $\begin{bmatrix} \frac{5}{3} & -\frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & \frac{5}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} & \frac{8}{3} \end{bmatrix}$

Ejercicios 5.5

1. $G' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, C' = C$

3. $G' = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, C' es equivalente a C, pero $C' \neq C$

5. $P' = [1 \quad 0 \quad 1]$, C' es equivalente a C, pero $C' \neq C$

7. $P' = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}$, C' es equivalente a C, pero $C' \neq C$

9. $C^\perp = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$

11. $C^\perp = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

13. $G^\perp = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, P^\perp = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$

15. $G^\perp = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}, P^\perp = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$

17. $G^\perp = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, P^\perp = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$

23. $2x^2 + 6xy + 4y^2$ **25.** 123

27. -5 **29.** $\begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$

31. $\begin{bmatrix} 3 & -\frac{3}{2} \\ -\frac{3}{2} & -1 \end{bmatrix}$ **33.** $\begin{bmatrix} 5 & 1 & -2 \\ 1 & -1 & 2 \\ -2 & 2 & 2 \end{bmatrix}$

35. $Q = \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ 1/\sqrt{5} & -2/\sqrt{5} \end{bmatrix}, y_1^2 + 6y_2^2$

37. $Q = \begin{bmatrix} 2/\sqrt{5} & 2/3\sqrt{5} & -1/3 \\ 0 & 5/3\sqrt{5} & 2/3 \\ 1/\sqrt{5} & -4/3\sqrt{5} & 2/3 \end{bmatrix}, 9y_1^2 + 9y_2^2 - 9y_3^2$

39. $Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ -1/\sqrt{3} & 0 & 2/\sqrt{6} \\ -1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \end{bmatrix}$, $2(x')^2 + (y')^2 - (z')^2$

41. Definida positiva **43.** Definida negativa

45. Definida positiva **47.** Indefinida

51. Para cualquier vector $\mathbf{x}$, se tiene $\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T B^T B\mathbf{x} = (B\mathbf{x})^T(B\mathbf{x}) = \|B\mathbf{x}\|^2 \geq 0$. Si $\mathbf{x}^T A\mathbf{x} = 0$, entonces $\|B\mathbf{x}\|^2 = 0$, de modo que $B\mathbf{x} = \mathbf{0}$. Puesto que B es invertible, esto implica que $\mathbf{x} = \mathbf{0}$. Por tanto $\mathbf{x}^T A\mathbf{x} > 0$ para todo $\mathbf{x} \neq \mathbf{0}$, y en consecuencia $A = B^T B$ es definida positiva.

53. **(a)** Todo eigenvalor de cA es de la forma $c\lambda$ para algún eigenvalor λ de A. Por el Teorema 5.24, $\lambda > 0$, de modo que $c\lambda > 0$, puesto que c es positivo. En consecuencia, cA es definida positiva, por el Teorema 5.24.

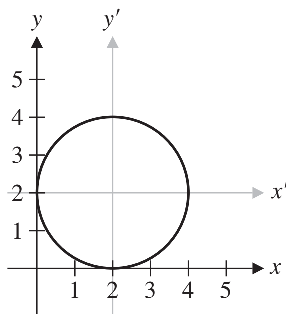
(c) Sea $\mathbf{x} \neq \mathbf{0}$. Entonces $\mathbf{x}^T A\mathbf{x} > 0$ y $\mathbf{x}^T B\mathbf{x} > 0$, puesto que A y B son definidas positivas. Pero entonces $\mathbf{x}^T(A + B)\mathbf{x} = \mathbf{x}^T A\mathbf{x} + \mathbf{x}^T B\mathbf{x} > 0$, de modo que $A + B$ es definida positiva.

55. El valor máximo de $f(\mathbf{x})$ es 2 cuando $\mathbf{x} = \pm\begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$; el valor mínimo de $f(\mathbf{x})$ es 0 cuando $\mathbf{x} = \pm\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

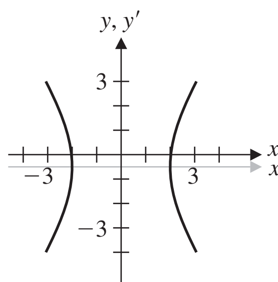
57. El valor máximo de $f(\mathbf{x})$ es 4 cuando $\mathbf{x} = \pm\begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$; el valor mínimo de $f(\mathbf{x})$ es 1 cuando $\mathbf{x} = \pm\begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$ o $\pm\begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$.

61. Elipse **63.** Parábola **65.** Hipérbola

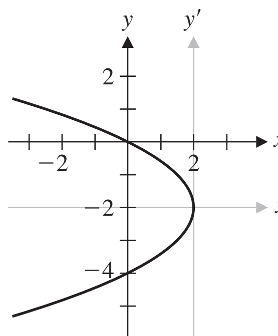
67. Circunferencia $x' = x - 2$, $y' = y - 2$, $(x')^2 + (y')^2 = 4$



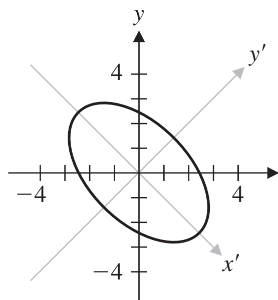
69. Hipérbola, $x' = x$, $y' = y + \frac{1}{2}$, $(x')^2/4 - (y')^2/9 = 1$



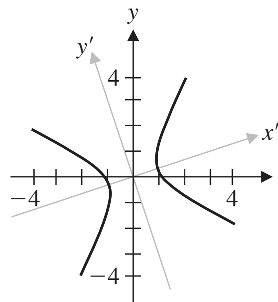
71. Parábola, $x' = x - 2$, $y' = y + 2$, $x' = -\frac{1}{2}(y')^2$



73. Elipse, $(x')^2/4 + (y')^2/12 = 1$



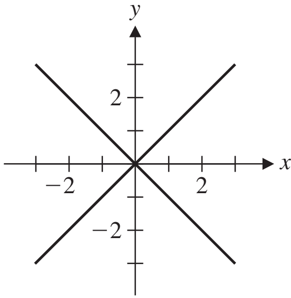
75. Hipérbola, $(x')^2 - (y')^2 = 1$



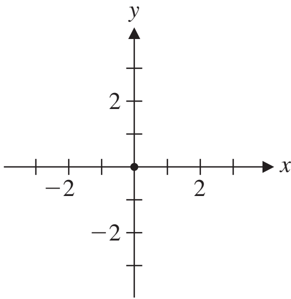
77. Elipse, $(x'')^2/50 + (y'')^2/10 = 1$

79. Hipérbola, $(x'')^2 - (y'')^2 = 1$

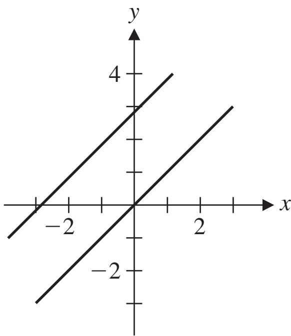
81. Degenerada (dos rectas)



83. Degenerada (un punto)



85. Degenerada (dos rectas)



89. Hiperboloide de una hoja, $(x')^2 - (y')^2 + 3(z')^2 = 1$

91. Paraboloide hiperbólico, $z = -(x')^2 + (y')^2$

93. Paraboloide hiperbólico, $x' = -\sqrt{3}(y')^2 + \sqrt{3}(z')^2$

95. Elipsoide, $3(x'')^2 + (y'')^2 + 2(z'')^2 = 4$

Preguntas de repaso

1. **(a)** V **(c)** V **(e)** F **(g)** F **(i)** F

3. $\begin{bmatrix} 9/2 \\ 2/3 \\ -11/6 \end{bmatrix}$

5. Verifique que $Q^TQ = I$.

7. El Teorema 5.6(c) muestra que si $\mathbf{v}_i \cdot \mathbf{v}_j = 0$, entonces $Q\mathbf{v}_i \cdot Q\mathbf{v}_j = 0$. El Teorema 5.6(b) muestra que $\{Q\mathbf{v}_1, \ldots, Q\mathbf{v}_k\}$ consiste de vectores unitarios, igual que $\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Por tanto, $\{Q\mathbf{v}_1, \ldots, Q\mathbf{v}_k\}$ es un conjunto ortonormal.

9. $\left\{\begin{bmatrix} 2 \\ -5 \end{bmatrix}\right\}$

11. $\left\{\begin{bmatrix} -1 \\ 3 \\ 1 \end{bmatrix}\right\}$

13. renglón(A): $\{[1 \quad 0 \quad 2 \quad 3 \quad 4], [0 \quad 1 \quad 0 \quad 2 \quad 1]\}$

col(A): $\left\{\begin{bmatrix} 1 \\ -1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 1 \\ -5 \end{bmatrix}\right\}$

nulo(A): $\left\{\begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -4 \\ -1 \\ 0 \\ 0 \\ 1 \end{bmatrix}\right\}$

nulo(A^T): $\left\{\begin{bmatrix} -5 \\ -3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 2 \\ 0 \\ 1 \end{bmatrix}\right\}$

15. **(a)** $\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ -\frac{3}{4} \end{bmatrix}, \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \end{bmatrix}\right\}$

17. $\left\{\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \\ -1 \end{bmatrix}\right\}$

19. $\begin{bmatrix} -\frac{1}{2} & \frac{3}{2} & 0 \\ \frac{3}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Capítulo 6

Ejercicios 6.1

1. Espacio vectorial

3. No es un espacio vectorial; falla el axioma 1.

5. No es un espacio vectorial; falla el axioma 8.

7. Espacio vectorial

9. Espacio vectorial

11. Espacio vectorial

15. Espacio vectorial complejo

17. No es un espacio vectorial complejo; falla el axioma 6.

19. No es un espacio vectorial; fallan los axiomas 1, 4 y 6.

21. No es un espacio vectorial; las operaciones de suma y multiplicación ni siquiera son las mismas.

25. Subespacio **27.** No es un subespacio

29. No es un subespacio **31.** Subespacio

33. Subespacio **35.** Subespacio

37. No es un subespacio **39.** Subespacio

41. Subespacio **43.** No es un subespacio

45. No es un subespacio

47. Tome U como el eje x y W como el eje y, por ejemplo. Entonces $\begin{bmatrix}1\\0\end{bmatrix}$ y $\begin{bmatrix}0\\1\end{bmatrix}$ están en $U \cup W$, pero $\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix}$ no lo está.

51. No

53. Sí; $s(x) = (3 + 2t)p(x) + (1 + t)q(x) + tr(x)$ para cualquier escalar t.

55. Sí; $h(x) = f(x) + g(x)$

57. No

59. No

61. Sí

Ejercicios 6.2

1. Linealmente independiente

3. Linealmente dependiente; $\begin{bmatrix}-1 & 0\\-1 & 7\end{bmatrix} = 4\begin{bmatrix}-1 & 1\\-2 & 2\end{bmatrix} + \begin{bmatrix}3 & 0\\1 & 1\end{bmatrix} - 2\begin{bmatrix}0 & 2\\-3 & 1\end{bmatrix}$

5. Linealmente independiente

7. Linealmente dependentiente; $3x + 2x^2 = 7x - 2(2x - x^2)$

9. Linealmente independiente

11. Linealmente dependiente; $1 = \operatorname{sen}^2 x + \cos^2 x$

13. Linealmente dependiente; $\ln(x^2) = -2\ln 2 \cdot 1 + 2 \cdot \ln(2x)$

17. **(a)** Linealmente independiente
(b) Linealmente dependiente

19. Base **21.** No es una base

23. No es una base **25.** No es una base

27. $[A]_{\mathcal{B}} = \begin{bmatrix}-1\\-1\\-1\\4\end{bmatrix}$ **29.** $[p(x)]_{\mathcal{B}} = \begin{bmatrix}6\\-1\\3\end{bmatrix}$

35. $\dim V = 2, \mathcal{B} = \{1 - x, 1 - x^2\}$

37. $\dim V = 3, \mathcal{B} = \left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right\}$

39. $\dim V = 2, \mathcal{B} = \left\{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\right\}$

41. $(n^2 - n)/2$

43. **(a)** $\dim(U \times V) = \dim U + \dim V$
(b) Demuestre que si $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ es una base para W, entonces $\{(\mathbf{w}_1, \mathbf{w}_1), \ldots, (\mathbf{w}_n, \mathbf{w}_n)\}$ es una base para Δ.

45. $\{1 + x, 1 + x + x^2, 1\}$

47. $\left\{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}, \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\right\}$

49. $\{1, 1 + x\}$

51. $\{1 - x, x - x^2\}$

53. $\{\operatorname{sen}^2 x, \cos^2 x\}$

59. **(a)** $p_0(x) = \frac{1}{2}x^2 - \frac{5}{2}x + 3, p_1(x) = -x^2 + 4x - 3, p_2(x) = \frac{1}{2}x^2 - \frac{3}{2}x + 1$

61. **(c)** (i) $3x^2 - 16x + 19$ (ii) $x^2 - 4x + 5$

63. $(p^n - 1)(p^n - p)(p^n - p^2)\cdots(p^n - p^{n-1})$

Ejercicios 6.3

1. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix}2\\3\end{bmatrix}, [\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix}\frac{5}{2}\\-\frac{1}{2}\end{bmatrix}, P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}\frac{1}{2} & \frac{1}{2}\\\frac{1}{2} & -\frac{1}{2}\end{bmatrix}, P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$

3. $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix}1\\0\\-1\end{bmatrix}, [\mathbf{x}]_{\mathcal{C}} = \begin{bmatrix}1\\-1\\-1\end{bmatrix}, P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & -1 & 1\end{bmatrix},$
$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix}1 & 0 & 0\\1 & 1 & 0\\1 & 1 & 1\end{bmatrix}$

5. $[p(x)]_{\mathcal{B}} = \begin{bmatrix}2\\-1\end{bmatrix}, [p(x)]_{\mathcal{C}} = \begin{bmatrix}-3\\2\end{bmatrix}, P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}-1 & 1\\1 & 0\end{bmatrix},$
$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix}0 & 1\\1 & 1\end{bmatrix}$

7. $[p(x)]_{\mathcal{B}} = \begin{bmatrix}1\\-1\\1\end{bmatrix}, [p(x)]_{\mathcal{C}} = \begin{bmatrix}1\\0\\1\end{bmatrix}, P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}1 & 0 & 0\\1 & 1 & 0\\1 & 1 & 1\end{bmatrix},$
$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix}1 & 0 & 0\\-1 & 1 & 0\\0 & -1 & 1\end{bmatrix}$

9. $[A]_{\mathcal{B}} = \begin{bmatrix} 4 \\ 2 \\ 0 \\ -1 \end{bmatrix}, [A]_{\mathcal{C}} = \begin{bmatrix} \frac{5}{2} \\ 0 \\ -3 \\ \frac{9}{2} \end{bmatrix}, P_{\mathcal{C}\leftarrow\mathcal{B}} =$

$\begin{bmatrix} \frac{1}{2} & 0 & -1 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 1 & 1 \\ \frac{3}{2} & -1 & -2 & -\frac{1}{2} \end{bmatrix}, P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 1 \end{bmatrix}$

11. $[f(x)]_{\mathcal{B}} = \begin{bmatrix} 2 \\ -5 \end{bmatrix}, [f(x)]_{\mathcal{C}} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, P_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix},$

$P_{\mathcal{B}\leftarrow\mathcal{C}} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}$

13. **(a)** $\begin{bmatrix} (3-2\sqrt{3})/2 \\ (-3\sqrt{3}+2)/2 \end{bmatrix} \approx \begin{bmatrix} 3.232 \\ -1.598 \end{bmatrix}$

(b) $\begin{bmatrix} 2+2\sqrt{3} \\ 2\sqrt{3}-2 \end{bmatrix} \approx \begin{bmatrix} 5.464 \\ 1.464 \end{bmatrix}$

15. $\mathcal{B} = \left\{ \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix} \right\}$

17. $-2 - 8(x-1) - 5(x-1)^2$

19. $-1 + 3(x+1) - 3(x+1)^2 + (x+1)^3$

Ejercicios 6.4

1. Transformación lineal **3.** Transformación lineal

5. Transformación lineal

7. No es una transformación lineal

9. Transformación lineal

11. No es una transformación lineal

13. Se tiene

$$\begin{aligned} S(p(x)+q(x)) &= S((p+q)(x)) = x((p+q)(x)) \\ &= x(p(x)+q(x)) = xp(x) + xq(x) \\ &= S(p(x)) + S(q(x)) \end{aligned}$$

y
$$\begin{aligned} S(cp(x)) &= S((cp)(x)) = x((cp)(x)) \\ &= x(cp(x)) = cxp(x) = cS(p(x)) \end{aligned}$$

Por tanto, S es lineal. De igual modo,

$$\begin{aligned} T\left(\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix}\right) &= T\begin{bmatrix} a+c \\ b+d \end{bmatrix} \\ &= (a+c) + ((a+c)+(b+d))x \\ &= (a+(a+b)x) + (c+(c+d)x) \\ &= T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) + T\left(\begin{bmatrix} c \\ d \end{bmatrix}\right) \end{aligned}$$

y
$$\begin{aligned} T\left(k\begin{bmatrix} a \\ b \end{bmatrix}\right) &= T\begin{bmatrix} ka \\ kb \end{bmatrix} = (ka) + (ka+kb)x \\ &= k(a+(a+b)x) = kT\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) \end{aligned}$$

Por tanto, T es lineal.

15. $T\begin{bmatrix} -7 \\ 9 \end{bmatrix} = 5 - 14x - 8x^2, T\begin{bmatrix} a \\ b \end{bmatrix} = \left(\dfrac{a+3b}{4}\right) - \left(\dfrac{a+7b}{4}\right)x + \left(\dfrac{a-b}{2}\right)x^2$

17. $T(4 - x + 3x^2) = 4 + 3x + 5x^2, T(a + bx + cx^2) = a + cx + \left(\dfrac{3a-b-c}{2}\right)x^2$

19. *Sugerencia:* sea $a = T(E_{11}), b = T(E_{12}), c = T(E_{21}), d = T(E_{22})$.

23. *Sugerencia:* considere el efecto de T y D sobre la base estándar para $\mathcal{P}_n$.

25. $(S \circ T)\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ 0 & 6 \end{bmatrix}, (S \circ T)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x & -y \\ 0 & 2x+2y \end{bmatrix}.$

$(T \circ S)\begin{bmatrix} x \\ y \end{bmatrix}$ no tiene sentido.

27. $(S \circ T)(p(x)) = p'(x+1), (T \circ S)(p(x)) = (p(x+1))' = p'(x+1)$

29. $(S \circ T)\begin{bmatrix} x \\ y \end{bmatrix} = S\left(T\begin{bmatrix} x \\ y \end{bmatrix}\right) = S\left(\begin{bmatrix} x-y \\ -3x+4y \end{bmatrix}\right) = \begin{bmatrix} 4(x-y)+(-3x+4y) \\ 3(x-y)+(-3x+4y) \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$

$(T \circ S)\begin{bmatrix} x \\ y \end{bmatrix} = T\left(S\begin{bmatrix} x \\ y \end{bmatrix}\right) = T\left(\begin{bmatrix} 4x+y \\ 3x+y \end{bmatrix}\right) = \begin{bmatrix} (4x+y)-(3x+y) \\ -3(4x+y)+4(3x+y) \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$

Por tanto, $S \circ T = I$ y $T \circ S = I$, de modo que S y T son inversos.

Ejercicios 6.5

1. **(a)** Sólo (ii) está en $\ker(T)$.
(b) Sólo (iii) está en el rango(T).
(c) $\ker(T) = \left\{\begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix}\right\}$, rango$(T) = \left\{\begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix}\right\}$

3. **(a)** Sólo (iii) está en $\ker(T)$.
(b) Todos ellos están en dominio(T).
(c) $\ker(T) = \{a + bx + cx^2 : a = -c, b = -c\} = \{t + tx - tx^2\}$, rango$(T) = \mathbb{R}^2$

5. Una base para ker(T) es $\left\{\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\right\}$ y una base para dominio (T) es $\left\{\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right\}$; rank($T$) = nulidad ($T$) = 2, y rank($T$) + nulidad ($T$) = 4 = dim M_{22}.

7. Una base para ker(T) es $\{1 + x - x^2\}$ y una base para dominio(T) es $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}$; rank($T$) = 2, nulidad) = 1, y rank(T) + nulidad (T) = 3 = dim $\mathscr{P}_2$.

9. rank(T) = nulidad(T) = 2

11. rank(T) = nulidad(T) = 2

13. rank(T) = 1, nulidad(T) = 2

15. Inyectiva y sobreyectiva

17. Ni inyectiva ni sobreyectiva

19. Inyectiva mas no sobreyectiva

21. Isomórfica, $T\begin{bmatrix}a & 0 & 0\\0 & b & 0\\0 & 0 & c\end{bmatrix} = \begin{bmatrix}a\\b\\c\end{bmatrix}$

23. No es isomórfica

25. Isomórfica, $T(a + bi) = \begin{bmatrix}a\\b\end{bmatrix}$

31. *Sugerencia:* defina $T: \mathscr{C}[0, 1] \to \mathscr{C}[0, 2]$ al hacer $T(f)$ la función cuyo valor en x es $(T(f))(x) = f(x/2)$ para x en $[0, 2]$.

33. (a) Sean $\mathbf{v}_1$ y $\mathbf{v}_2$ que están en V y sea $(S \circ T)(\mathbf{v}_1) = (S \circ T)(\mathbf{v}_2)$. Entonces $S(T(\mathbf{v}_1)) = S(T(\mathbf{v}_2))$, por tanto $T(\mathbf{v}_1) = T(\mathbf{v}_2)$, puesto que S es inyectiva. Pero ahora $\mathbf{v}_1 = \mathbf{v}_2$, pues T es inyectiva. En consecuencia, $S \circ T$ es inyectiva.

35. (a) Por el teorema del rank, rank(T) + nulidad(T) = dim V. Si T es sobreyectiva, entonces rango(T) = W, de modo que rank(T) = dim(rank(T)) = dim W. Por tanto, dim V + nulidad(T) < dim W + nulidad(T) = rank(T) + nulidad(T) = dim V de modo que nulidad(T) < 0, lo cual es imposible. Por tanto, T no puede ser sobreyectiva.

Ejercicios 6.6

1. $[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}$, $[T]_{\mathcal{C}\leftarrow\mathcal{B}}[4 + 2x]_{\mathcal{B}} = \begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}\begin{bmatrix}4\\2\end{bmatrix} = \begin{bmatrix}2\\-4\end{bmatrix} = [2 - 4x]_{\mathcal{C}} = [T(4 + 2x)]_{\mathcal{C}}$

3. $[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$, $[T]_{\mathcal{C}\leftarrow\mathcal{B}}[a + bx + cx^2]_{\mathcal{B}} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}a\\b\\c\end{bmatrix} = [a + b(x + 2) + c(x + 2)^2)]_{\mathcal{C}} = [T(a + bx + cx^2)]_{\mathcal{C}}$

5. $[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}1 & 0 & 0\\1 & 1 & 1\end{bmatrix}$, $[T]_{\mathcal{C}\leftarrow\mathcal{B}}[a + bx + cx^2]_{\mathcal{B}} = \begin{bmatrix}1 & 0 & 0\\1 & 1 & 1\end{bmatrix}\begin{bmatrix}a\\b\\c\end{bmatrix} = \begin{bmatrix}a\\a + b + c\end{bmatrix} = \begin{bmatrix}a + b\cdot 0 + c\cdot 0^2\\a + b\cdot 1 + c\cdot 1^2\end{bmatrix}_{\mathcal{C}} = [T(a + bx + cx^2)]_{\mathcal{C}}$

7. $[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}6 & 4\\-3 & -2\\2 & -1\end{bmatrix}$, $[T]_{\mathcal{C}\leftarrow\mathcal{B}}\begin{bmatrix}-7\\7\end{bmatrix}_{\mathcal{B}} = \begin{bmatrix}6 & 4\\-3 & -2\\2 & -1\end{bmatrix}\begin{bmatrix}2\\-3\end{bmatrix} = \begin{bmatrix}0\\0\\7\end{bmatrix} = \begin{bmatrix}7\\7\\7\end{bmatrix}_{\mathcal{C}} = \begin{bmatrix}7\\7\\7\end{bmatrix}_{\mathcal{C}} = \left[T\begin{bmatrix}-7\\7\end{bmatrix}\right]_{\mathcal{C}}$

9. $[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}1 & 0 & 0 & 0\\0 & 0 & 1 & 0\\0 & 1 & 0 & 0\\0 & 0 & 0 & 1\end{bmatrix}$, $[T]_{\mathcal{C}\leftarrow\mathcal{B}}[A]_{\mathcal{B}} = \begin{bmatrix}1 & 0 & 0 & 0\\0 & 0 & 1 & 0\\0 & 1 & 0 & 0\\0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}a\\b\\c\\d\end{bmatrix} = \begin{bmatrix}a\\c\\b\\d\end{bmatrix} = \left[\begin{bmatrix}a & c\\b & d\end{bmatrix}\right]_{\mathcal{C}} = [T(A)]_{\mathcal{C}}$

11. $[T]_{\mathcal{C}\leftarrow\mathcal{B}} = \begin{bmatrix}0 & -1 & 1 & 0\\-1 & 0 & 0 & 1\\1 & 0 & 0 & -1\\0 & 1 & -1 & 0\end{bmatrix}$, $[T]_{\mathcal{C}\leftarrow\mathcal{B}}[A]_{\mathcal{B}} = \begin{bmatrix}0 & -1 & 1 & 0\\-1 & 0 & 0 & 1\\1 & 0 & 0 & -1\\0 & 1 & -1 & 0\end{bmatrix}\begin{bmatrix}a\\b\\c\\d\end{bmatrix} = \begin{bmatrix}c - b\\d - a\\a - d\\b - c\end{bmatrix} = \left[\begin{bmatrix}c - b & d - a\\a - d & b - c\end{bmatrix}\right]_{\mathcal{C}} = [AB - BA]_{\mathcal{C}} = [T(A)]_{\mathcal{C}}$

13. (b) $[D]_{\mathcal{B}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

(c) $[D]_{\mathcal{B}}[3 \text{ sen } x - 5 \cos x]_{\mathcal{B}} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 3 \\ -5 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} = [3 \cos x + 5 \text{ sen } x]_{\mathcal{B}} = [D(3 \text{ sen } x - 5 \cos x)]_{\mathcal{B}}$

15. (a) $[D]_{\mathcal{B}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & -1 & 2 \end{bmatrix}$

17. $[S \circ T]_{\mathcal{D} \leftarrow \mathcal{B}} = \begin{bmatrix} -1 & -2 \\ 1 & -1 \end{bmatrix}$

19. Invertible, $T^{-1}(a + bx) = -b + ax$

21. Invertible, $T^{-1}(p(x)) = p(x - 2)$

23. Invertible, $T^{-1}(a + bx + cx^2) = (a - b + 2c) + (b - 2c)x + cx^2$ o $T^{-1}(p(x)) = p(x) - p'(x) + p''(x)$

25. No es invertible **27.** $-3 \text{ sen } x - \cos x + C$

29. $\frac{4}{5}e^{2x} \cos x - \frac{3}{5}e^{2x} \text{ sen } x + C$

31. $\mathcal{C} = \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \end{bmatrix} \right\}$ **33.** $\mathcal{C} = \{1 - x, 2 + x\}$

35. $\mathcal{C} = \{1, x\}$

37. $[T]_{\mathcal{E}} = \begin{bmatrix} (d_1^2 - d_2^2)/(d_1^2 + d_2^2) & 2d_1d_2/(d_1^2 + d_2^2) \\ 2d_1d_2/(d_1^2 + d_2^2) & (d_2^2 - d_1^2)/(d_1^2 + d_2^2) \end{bmatrix}$

Ejercicios 6.7

1. $y(t) = 2e^{3t}/e^3$

3. $y(t) = ((1 - e^4)e^{3t} + (e^3 - 1)e^{4t})/(e^3 - e^4)$

5. $f(t) = \left(\dfrac{e^{(\sqrt{5}-1)/2}}{e^{\sqrt{5}} - 1} \right)[e^{(1+\sqrt{5})t/2} - e^{(1-\sqrt{5})t/2}]$

7. $y(t) = e^t - (1 - e^{-1})te^t$

9. $y(t) = ((k + 1)e^{kt} + (k - 1)e^{-kt})/2k$

11. $y(t) = e^t \cos(2t)$

13. (a) $p(t) = 100e^{\ln(16)t/3} \approx 100e^{0.924t}$
(b) 45 minutos **(c)** En 9.968 horas

15. (a) $m(t) = 50e^{-ct}$, donde $c = \ln 2/1590 \approx 4.36 \times 10^{-4}$; después de 1000 años quedan 32.33 mg
(b) Después de 3691.9 años

17. $x(t) = \dfrac{5 - 10 \cos(10\sqrt{K})}{\text{sen}(10\sqrt{K})} \text{sen}(\sqrt{K}t) + 10 \cos(\sqrt{K}t)$

19. (b) No **23.** No es lineal

25. No es lineal **27.** Lineal

29. Lineal

31. $\left\{ \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}, \left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$

35. $R_3 = \left\{ \begin{bmatrix} 0\\0\\0\\0\\0\\0\\0\\0 \end{bmatrix} \begin{bmatrix} 0\\0\\0\\0\\1\\1\\1\\1 \end{bmatrix} \begin{bmatrix} 0\\0\\1\\1\\0\\0\\1\\1 \end{bmatrix} \begin{bmatrix} 0\\0\\1\\1\\1\\1\\0\\0 \end{bmatrix} \begin{bmatrix} 0\\1\\0\\1\\0\\1\\0\\1 \end{bmatrix} \begin{bmatrix} 0\\1\\0\\1\\1\\0\\1\\0 \end{bmatrix} \begin{bmatrix} 0\\1\\1\\0\\0\\1\\1\\0 \end{bmatrix} \begin{bmatrix} 0\\1\\1\\0\\1\\0\\0\\1 \end{bmatrix} \begin{bmatrix} 1\\0\\1\\0\\1\\0\\1\\0 \end{bmatrix} \begin{bmatrix} 1\\0\\1\\0\\0\\1\\0\\1 \end{bmatrix} \begin{bmatrix} 1\\0\\0\\1\\1\\0\\0\\1 \end{bmatrix} \begin{bmatrix} 1\\0\\0\\1\\0\\1\\1\\0 \end{bmatrix} \begin{bmatrix} 1\\1\\0\\0\\1\\1\\0\\0 \end{bmatrix} \begin{bmatrix} 1\\1\\0\\0\\0\\0\\1\\1 \end{bmatrix} \begin{bmatrix} 1\\1\\1\\1\\0\\0\\0\\0 \end{bmatrix} \begin{bmatrix} 1\\1\\1\\1\\1\\1\\1\\1 \end{bmatrix} \right\}$

37. $[1 \quad 1 \quad 1 \quad 1]$

Preguntas de repaso

1. (a) F **(c)** V **(e)** F **(g)** F **(i)** V

3. Subespacio **5.** Subespacio

7. Sea $c_1A + c_2B = O$. Entonces $c_1A - c_2B = c_1A^T + c_2B^T = (c_1A + c_2B)^T = O$. Al sumar, se tiene $2c_1A = O$, de modo que $c_1 = 0$ puesto que A es distinto de cero. Por tanto $c_2B = O$, y en consecuencia $c_2 = 0$. Por ende, $\{A, B\}$ es linealmente independiente.

9. $\{1, x^2, x^4\}$, dim $W = 3$

11. Transformación lineal

13. Transformación lineal **15.** $n^2 - 1$

17. $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}$

19. $S \circ T$ es la transformación cero.

Capítulo 7

Ejercicios 7.1

1. (a) 0 **(b)** $\sqrt{11}$ **(c)** $\sqrt{77}$

3. Cualquier múltiplo escalar distinto de cero de $\mathbf{v} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

5. (a) -1 **(b)** $\sqrt{14}$ **(c)** $\sqrt{26}$

7. $1 - 2x^2$

9. (a) π **(b)** $\sqrt{\pi}$ **(c)** $\sqrt{\pi}$

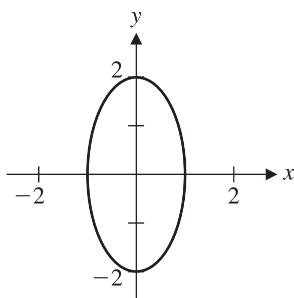
13. Falla el axioma (4): $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq \mathbf{0}$, pero $\langle \mathbf{u}, \mathbf{u} \rangle = 0$.

15. Falla el axioma (4): $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq \mathbf{0}$, pero $\langle \mathbf{u}, \mathbf{u} \rangle = 0$.

17. Falla el axioma (4): $p(x) = 1 - x$ no es el polinomio cero, pero $\langle p(x), p(x)\rangle = 0$.

19. $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$

21.



25. -8 **27.** $\sqrt{6}$

29. $\|\mathbf{u} + \mathbf{v} - \mathbf{w}\|^2 = \langle \mathbf{u} + \mathbf{v} - \mathbf{w}, \mathbf{u} + \mathbf{v} - \mathbf{w}\rangle$

$$= \langle \mathbf{u}, \mathbf{u}\rangle + \langle \mathbf{v}, \mathbf{v}\rangle + \langle \mathbf{w}, \mathbf{w}\rangle + 2\langle \mathbf{u}, \mathbf{v}\rangle - 2\langle \mathbf{u}, \mathbf{w}\rangle - 2\langle \mathbf{v}, \mathbf{w}\rangle$$
$$= 1 + 3 + 4 + 2 - 10 - 0 = 0$$

Por tanto, $\|\mathbf{u} + \mathbf{v} - \mathbf{w}\| = 0$, de modo que, por el axioma (4), $\mathbf{u} + \mathbf{v} - \mathbf{w} = \mathbf{0}$ o $\mathbf{u} + \mathbf{v} = \mathbf{w}$.

31. $\langle \mathbf{u} + \mathbf{v}, \mathbf{u} - \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{u}\rangle - \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{v}, \mathbf{u}\rangle - \langle \mathbf{v}, \mathbf{v}\rangle = \|\mathbf{u}\|^2 - \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{v}\rangle - \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2$

33. Al usar el ejercicio 32 y una identidad similar para $\|\mathbf{u} - \mathbf{v}\|^2$, se tiene

$$\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v}\rangle + \langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v}\rangle$$
$$= \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 + \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2$$
$$= 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

Al dividir por 2 produce la identidad que se quiere.

35. $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u} - \mathbf{v}\| \Leftrightarrow \|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u} - \mathbf{v}\|^2$

$$\Leftrightarrow \|\mathbf{u}\|^2 + 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 - 2\langle \mathbf{u}, \mathbf{v}\rangle + \|\mathbf{v}\|^2$$
$$\Leftrightarrow 2\langle \mathbf{u}, \mathbf{v}\rangle = -2\langle \mathbf{u}, \mathbf{v}\rangle \Leftrightarrow \langle \mathbf{u}, \mathbf{v}\rangle = 0$$

37. $\left\{\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right\}$ **39.** $\{1, x, x^2\}$

41. **(a)** $1/\sqrt{2}, \sqrt{3}x/\sqrt{2}, \sqrt{5}(3x^2 - 1)/2\sqrt{2}$
(b) $\sqrt{7}(5x^3 - 3x)/2\sqrt{2}$

Ejercicios 7.2

1. $\|\mathbf{u}\|_E = \sqrt{42}$, $\|\mathbf{u}\|_s = 10$, $\|\mathbf{u}\|_m = 5$

3. $d_E(\mathbf{u}, \mathbf{v}) = \sqrt{70}$, $d_s(\mathbf{u}, \mathbf{v}) = 14$, $d_m(\mathbf{u}, \mathbf{v}) = 6$

5. $\|\mathbf{u}\|_H = 4$, $\|\mathbf{v}\|_H = 5$

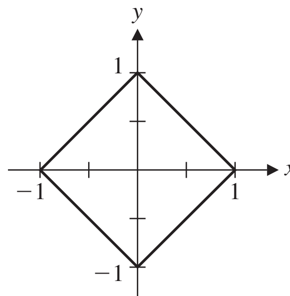
7. **(a)** Como máximo un componente de $\mathbf{v}$ es distinto de cero.

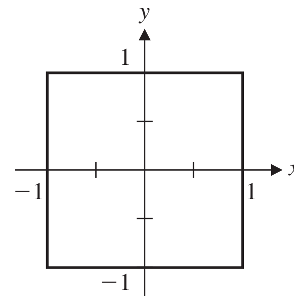
9. Suponga $\|\mathbf{v}\|_m = |v_k|$. Entonces $\|\mathbf{v}\|_E = \sqrt{v_1^2 + \cdots + v_k^2 + \cdots + v_n^2} \geq \sqrt{v_k^2} = |v_k| = \|\mathbf{v}\|_m$.

11. Suponga $\|\mathbf{v}\|_m = |v_k|$. Entonces $|v_i| \leq |v_k|$ para $i = 1, \ldots, n$, de modo que

$$\|\mathbf{v}\|_s = |v_1| + \cdots + |v_n| \leq |v_k| + \cdots + |v_k| = n|v_k| = n\|\mathbf{v}\|_m$$

13.





21. $\|A\|_F = \sqrt{19}$, $\|A\|_1 = 4$, $\|A\|_\infty = 6$

23. $\|A\|_F = \sqrt{31}$, $\|A\|_1 = 6$, $\|A\|_\infty = 6$

25. $\|A\|_F = 2\sqrt{11}$, $\|A\|_1 = 7$, $\|A\|_\infty = 7$

27. $\mathbf{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ **29.** $\mathbf{x} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

31. $\mathbf{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$

33. **(a)** Por la definición de operador norma, $\|I\| = \max_{\|\mathbf{x}\|=1} \|I\mathbf{x}\| = \max_{\|\mathbf{x}\|=1} \|\mathbf{x}\| = 1$.

35. $\text{cond}_1(A) = \text{cond}_\infty(A) = 21$; bien condicionada

37. $\text{cond}_1(A) = \text{cond}_\infty(A) = 400$; mal condicionada

39. $\text{cond}_1(A) = 77$, $\text{cond}_\infty(A) = 128$; moderadamente mal condicionada

41. **(a)** $\text{cond}_\infty(A) = (\text{máx}\{|k| + 1, 2\}) \cdot \left(\text{máx}\left\{\left|\frac{k}{k-1}\right| + \left|\frac{1}{k-1}\right|, \left|\frac{2}{k-1}\right|\right\}\right)$

43. **(a)** $\text{cond}_\infty(A) = 40$
(b) Como máximo 400% de cambio relativo

45. Al usar el ejercicio 33(a) se tiene $\text{cond}(A) = \|A\|\|A^{-1}\| \geq \|AA^{-1}\| = \|I\| = 1$.

49. $k \geq 6$ **51.** $k \geq 10$

Ejercicios 7.3

1. $\|\mathbf{e}\| = \sqrt{2} \approx 1.414$ **3.** $\|\mathbf{e}\| = \sqrt{6}/2 \approx 1.225$

5. $\|\mathbf{e}\| = \sqrt{7} \approx 2.646$

7. $y = -3 + \frac{5}{2}x$, $\|\mathbf{e}\| \approx 1.225$

9. $y = \frac{11}{3} - 2x$, $\|\mathbf{e}\| \approx 0.816$

11. $y = \frac{7}{10} + \frac{8}{25}x$, $\|\mathbf{e}\| \approx 0.447$

13. $y = -\frac{1}{5} + \frac{7}{5}x$, $\|\mathbf{e}\| \approx 0.632$

15. $y = 3 - \frac{18}{5}x + x^2$ **17.** $y = \frac{18}{5} - \frac{17}{10}x - \frac{1}{2}x^2$

19. $\overline{\mathbf{x}} = \begin{bmatrix} \frac{1}{5} \\ \frac{7}{15} \end{bmatrix}$ **21.** $\overline{\mathbf{x}} = \begin{bmatrix} \frac{4}{3} \\ -\frac{5}{6} \end{bmatrix}$

23. $\overline{\mathbf{x}} = \begin{bmatrix} 4+t \\ -5-t \\ -5-2t \\ t \end{bmatrix}$ **25.** $\begin{bmatrix} \frac{42}{11} \\ \frac{19}{11} \\ \frac{42}{11} \end{bmatrix}$

27. $\overline{\mathbf{x}} = \begin{bmatrix} \frac{5}{3} \\ -2 \end{bmatrix}$ **29.** $y = 0.92 + 0.73x$

31. **(a)** Si se hace que el año 1920 corresponda a $t = 0$, entonces $y = 56.6 + 2.9t$; 79.9 años

33. **(a)** $p(t) = 150e^{0.131t}$

35. 139 días

37. $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}, \begin{bmatrix} \frac{7}{2} \\ \frac{7}{2} \end{bmatrix}$ **39.** $\begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix}$

41. $\begin{bmatrix} \frac{5}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & \frac{1}{3} & \frac{5}{6} \end{bmatrix}, \begin{bmatrix} \frac{5}{6} \\ \frac{1}{3} \\ -\frac{1}{6} \end{bmatrix}$ **45.** $A^+ = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \end{bmatrix}$

47. $A^+ = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & -\frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}$ **49.** $A^+ = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$

51. $A^+ = \begin{bmatrix} \frac{2}{3} & 0 & -\frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -1 & \frac{1}{3} & \frac{2}{3} \\ -\frac{2}{3} & 1 & \frac{1}{3} & -\frac{1}{3} \end{bmatrix}$

53. **(a)** Si A es invertible, también lo es A^T, y se tiene $A^+ = (A^TA)^{-1}A^T = A^{-1}(A^T)^{-1}A^T = A^{-1}$.

Ejercicios 7.4

1. 2, 3 **3.** $\sqrt{2}, 0$ **5.** 5 **7.** 2, 3

9. $\sqrt{5}, 2, 0$

11. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$

13. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

15. $A = \begin{bmatrix} \frac{3}{5} & -\frac{4}{5} \\ \frac{4}{5} & \frac{3}{5} \end{bmatrix} \begin{bmatrix} 5 \\ 0 \end{bmatrix} [1]$

17. $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

19. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{5} & 0 & 0 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 0 & 1/\sqrt{5} \\ 0 & 1 & 0 \\ 1/\sqrt{5} & 0 & -2/\sqrt{5} \end{bmatrix}$

21. $A = \sqrt{2}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} + 0\begin{bmatrix} 0 \\ 1 \end{bmatrix}$
$\begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ (Ejercicio 3)

23. (Ejercicio 7) $A = 3\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} + 2\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}$

33. El segmento de recta $[-1, 1]$

35. La elipse sólida $\frac{y_1^2}{5} + \frac{y_2^2}{4} \leq 1$

37. **(a)** $\|A\|_2 = \sqrt{2}$ **(b)** $\text{cond}_2(A) = \infty$

39. **(a)** $\|A\|_2 = 1.95$ **(b)** $\text{cond}_2(A) = 38.11$

41. $A^+ = \begin{bmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & 0 \end{bmatrix}$ **43.** $A^+ = \begin{bmatrix} \frac{2}{5} & 0 \\ 0 & \frac{1}{2} \\ \frac{1}{5} & 0 \end{bmatrix}$

45. $A^+ = \begin{bmatrix} \frac{1}{25} & \frac{2}{25} \\ \frac{2}{25} & \frac{4}{25} \end{bmatrix}, \overline{\mathbf{x}} = \begin{bmatrix} 0.52 \\ 1.04 \end{bmatrix}$

47. $A^+ = \begin{bmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \end{bmatrix}, \overline{\mathbf{x}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

61. $\begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

63. $\begin{bmatrix} 2 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

Ejercicios 7.5

1. $g(x) = \frac{1}{3}$ **3.** $g(x) = \frac{3}{5}x$

5. $g(x) = \frac{3}{16} + \frac{15}{16}x^2$ **7.** $\{1, x - \frac{1}{2}\}$

9. $g(x) = x - \frac{1}{6}$

11. $g(x) = (4e - 10) + (18 - 6e)x \approx 0.87 + 1.69x$

13. $g(x) = \frac{1}{20} - \frac{3}{5}x + \frac{3}{2}x^2$

15. $g(x) = 39e - 105 + (588 - 216e)x + (210e - 570)x^2 \approx 1.01 + 0.85x + 0.84x^2$

21. $\frac{\pi}{2} - \frac{4}{\pi}\left(\cos x + \frac{\cos 3x}{9}\right)$

23. $a_0 = \frac{1}{2}, a_k = 0, b_k = \frac{1 - (-1)^k}{k\pi}$

25. $a_0 = \pi, a_k = 0, b_k = \frac{2(-1)^k}{k}$

29. $d(C) = 1$ **31.** $d(C) = 2$ **33.** $d(C) = 3$

35. $d(C) = 3$; si $C = \{\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4\}$, entonces **u** se decodifica como $\mathbf{c}_2$, **v** no puede decodificarse y **w** se decodifica como $\mathbf{c}_3$.

37. $C = \{\mathbf{0}, \mathbf{1}\}$, donde **0** es el vector cero y **1** es el vector de todos los números 1 en $\mathbb{Z}_2^8$.

39. Una matriz de control de paridad P para tal código es $(8 - 5) \times 8 = 3 \times 8$. Por tanto, $\text{rank}(P) \leq 3$, de modo que cualesquiera cuatro columnas de P deben ser linealmente dependientes. Por tanto, el entero *más pequeño d* para el cual existen d columnas linealmente dependientes satisface $d \leq 4$. Por el Teorema 7.21, esto significa que $d(C) \leq 4$, de modo que no existe un código lineal $(8, 5, 5)$.

Preguntas de repaso

1. **(a)** V **(c)** F **(e)** V **(g)** V **(i)** V

3. Producto interno **5.** $\sqrt{3}$ **7.** $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}, \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$

9. No es una norma **11.** $\text{cond}_\infty(A) \approx 2432$

13. $y = 1.7x$ **15.** $\begin{bmatrix} \frac{7}{3} \\ \frac{2}{3} \\ \frac{5}{3} \end{bmatrix}$

17. **(a)** $\sqrt{2}, \sqrt{2}$

(b) $A = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \end{bmatrix}\begin{bmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(c) $A^+ = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix}$

19. Los valores singulares de PAQ son las raíces cuadradas de los eigenvalores de $(PAQ)^T(PAQ) = Q^TA^TP^TPAQ = Q^T(A^TA)Q$. Pero $Q^T(A^TA)Q$ es similar a A^TA porque $Q^T = Q^{-1}$, y en consecuencia tiene los mismos eigenvalores que A^TA. Por tanto, PAQ y A tienen los mismos valores singulares.

Índice

I've got a little list.
—Sir W. S. Gilbert, "They'll None Of 'Em Be Missed," om *The Mikado,* 1885

S

T

U

V

W

Z

EXPLORACIONES Y VIÑETAS

APLICACIONES

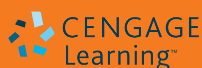
Índice de notación



La innovadora obra de David Poole destaca vectores e intuición geométrica desde el principio y prepara mejor al estudiante para hacer la transición de los aspectos computacionales del curso a los teóricos. Diseñado para un curso de introducción de uno o dos semestres y escrito de manera sencilla, "el matemático inglés" Poole centra su enfoque sobre la visualización que beneficia al estudiante y su conexión con el material. Ofrece ejemplos concretos para introducir al estudiante antes de presentar la abstracción y sigue inmediatamente una discusión teórica con ejemplos adicionales y una matriz de aplicaciones a una gran variedad de disciplinas. Estudiantes de una diversidad de fondos y estilos de aprendizaje se benefician de enfoque práctico de Poole, que cubre vectores y geometría vectorial desde el inicio con el fin de permitir que los estudiantes visualicen las matemáticas mientras están realizando operaciones matriciales. Con una comprensión concreta de la geometría de vectores, los estudiantes son capaces de visualizar y comprender el significado de los cálculos que se encuentran y desarrollar la madurez matemática para pensar en forma abstracta.

Características

- Exploraciones (1 por capítulo) proporcionan guías más profundas para el descubrimiento basadas sobre los conceptos claves, diseñados para trabajo individual o de grupo.
- Más de 400 ejemplos, generalmente trabajados con mayor detalle y más énfasis en la legibilidad que la mayoría de los libros.
- Los más de 2000 ejercicios y problemas aplicados a partir de una amplia variedad de disciplinas de ingeniería, ciencias físicas, ciencias biológicas y empresariales.
- Notas al margen sensibles al contexto para ayuda adicional y referencias cruzadas.
- Viñetas destacando aplicaciones del mundo real en los negocios, la ciencia y la sociedad con debate ampliado de los conceptos detrás de las aplicaciones.



http://latinoamerica.cengage.com



ISBN-13: 978-607481748-5
ISBN-10: 607481748-0

